# Inside: PLT Scheme C API

Version 4.0.1

June 22, 2008

This manual describes PLT Scheme's C interface, which allows the interpreter to be extended by a dynamically-loaded library, or embedded within an arbitrary C/C++ program. The manual assumes familiarity with PLT Scheme as described in §"**Reference**: PLT Scheme".

For an alternative way of dealing with foreign code, see §"FFI: PLT Scheme Foreign Interface", which describes the scheme/foreign module for manipulating low-level libraries and structures purely through Scheme code.

# Contents

| 1 | Ove  | rview                                      | 5  |
|---|------|--|----|
|   | 1.1  | CGC versus 3m                              | 5  |
|   | 1.2  | Writing MzScheme Extensions                | 5  |
|   |      | 1.2.1 CGC Extensions                       | 5  |
|   |      | 1.2.2 3m Extensions                        | 7  |
|   | 1.3  | Embedding MzScheme into a Program          | 8  |
|   |      | 1.3.1 CGC Embedding                        | 8  |
|   |      | 1.3.2 3m Embedding                         | 11 |
|   | 1.4  | MzScheme and Threads                       | 14 |
|   | 1.5  | MzScheme, Unicode, Characters, and Strings | 14 |
|   | 1.6  | Integers                                   | 14 |
| 2 | Valu | nes and Types                              | 15 |
|   | 2.1  | Standard Types                             | 16 |
|   | 2.2  | Global Constants                           | 19 |
|   | 2.3  | Strings                                    | 19 |
|   | 2.4  | Value Functions                            | 19 |
| 3 | Men  | nory Allocation                            | 29 |
|   | 3.1  | Cooperating with 3m                        | 30 |
|   |      | 3.1.1 Tagged Objects                       | 31 |
|   |      | 3.1.2 Local Pointers                       | 32 |
|   |      | 3.1.3 Local Pointers and mzcxform          | 36 |
|   |      | 3.1.4 Guiding mzcxform                     | 37 |
|   | 3.2  | Memory Functions                           | 39 |

| 4  | Nam  | nespaces and Modules                        | 46 |
|----|------|---|----|
| 5  | Proc | redures                                     | 49 |
| 6  | Eval | uation                                      | 52 |
|    | 6.1  | Top-level Evaluation Functions              | 52 |
|    | 6.2  | Tail Evaluation                             | 52 |
|    | 6.3  | Multiple Values                             | 53 |
|    | 6.4  | Evaluation Functions                        | 54 |
| 7  | Exce | eptions and Escape Continuations            | 58 |
|    | 7.1  | Temporarily Catching Error Escapes          | 59 |
|    | 7.2  | Enabling and Disabling Breaks               | 61 |
|    | 7.3  | Exception Functions                         | 62 |
| 8  | Thre | eads  | 67 |
|    | 8.1  | Integration with Threads                    | 67 |
|    | 8.2  | Allowing Thread Switches                    | 68 |
|    | 8.3  | Blocking the Current Thread                 | 68 |
|    | 8.4  | Threads in Embedded Scheme with Event Loops | 69 |
|    |      | 8.4.1 Callbacks for Blocked Threads         | 70 |
|    | 8.5  | Sleeping by Embedded Scheme                 | 72 |
|    | 8.6  | Thread Functions                            | 73 |
| 9  | Para | nmeterizations                              | 81 |
| 10 | Con  | tinuation Marks                             | 85 |
| 11 | Stri | ng Encodings                                | 86 |

| 12  | Bignums, Rationals, and Complex Numbers | 91  |
|-----|---|-----|
| 13  | Ports and the Filesystem                | 95  |
| 14  | Structures                              | 113 |
| 15  | Security Guards                         | 116 |
| 16  | Custodians                              | 118 |
| 17  | Miscellaneous Utilities                 | 120 |
| 18  | Flags and Hooks                         | 127 |
| Ind | lev                                     | 129 |

# 1 Overview

#### 1.1 CGC versus 3m

Before mixing any C code with MzScheme, first decide whether to use the **3m** variant of PLT Scheme, the **CGC** variant of PLT Scheme, or both:

- **3m**: the main variant of PLT Scheme, which uses *precise* garbage collection instead of conservative garbage collection, and it may move objects in memory during a collection.
- **CGC**: the original variant of PLT Scheme, where memory management depends on a *conservative* garbage collector. The conservative garbage collector can automatically find references to managed values from C local variables and (on some platforms) static variables.

At the C level, working with CGC can be much easier than working with 3m, but overall system performance is typically better with 3m.

# 1.2 Writing MzScheme Extensions

The process of creating an extension for 3m or CGC is essentially the same, but the process for 3m is most easily understood as a variant of the process for CGC.

#### 1.2.1 CGC Extensions

To write a C/C++-based extension for PLT Scheme CGC, follow these steps:

- For each C/C++ file that uses PLT Scheme library functions, #include the file "escheme.h".
  - This file is distributed with the PLT software in an "include" directory, but if mzc is used to compile, this path is found automatically.
- Define the C function scheme\_initialize, which takes a Scheme\_Env\* namespace (see §4 "Namespaces and Modules") and returns a Scheme\_Object\* Scheme value. This initialization function can install new global primitive procedures or other values into the namespace, or it can simply return a Scheme value. The initialization function is called when the extension is loaded with load-extension (the first time); the return value from scheme\_initialize is used as the return value for load-extension. The namespace provided to scheme\_initialize is the current namespace when load-extension is called.

 Define the C function scheme\_reload, which has the same arguments and return type as scheme\_initialize.

This function is called if load-extension is called a second time (or more times) for an extension. Like scheme\_initialize, the return value from this function is the return value for load-extension.

 Define the C function scheme\_module\_name, which takes no arguments and returns a Scheme\_Object\* value, either a symbol or scheme\_false.

The function should return a symbol when the effect of calling scheme\_initialize and scheme\_reload is only to declare a module with the returned name. This function is called when the extension is loaded to satisfy a require declaration.

The scheme\_module\_name function may be called before scheme\_initialize and scheme\_reload, after those functions, or both before and after, depending on how the extension is loaded and re-loaded.

• Compile the extension C/C++ files to create platform-specific object files.

The mzc compiler, which is distributed with PLT Scheme, compiles plain C files when the --cc flag is specified. More precisely, mzc does not compile the files itself, but it locates a C compiler on the system and launches it with the appropriate compilation flags. If the platform is a relatively standard Unix system, a Windows system with either Microsoft's C compiler or gcc in the path, or a Mac OS X system with Apple's developer tools installed, then using mzc is typically easier than working with the C compiler directly. Use the --cgc flag to indicate that the build is for use with PLT Scheme CGC.

• Link the extension C/C++ files with "mzdyn.o" (Unix, Mac OS X) or "mzdyn.obj" (Windows) to create a shared object. The resulting shared object should use the extension ".so" (Unix), ".dll" (Windows), or ".dylib" (Mac OS X).

The "mzdyn" object file is distributed in the installation's "lib" directory. For Windows, the object file is in a compiler-specific sub-directory of "plt\lib".

The mzc compiler links object files into an extension when the --ld flag is specified, automatically locating "mzdyn". Again, use the --cgc flag with mzc.

• Load the shared object within Scheme using (load-extension path), where path is the name of the extension file generated in the previous step.

Alternately, if the extension defines a module (i.e., scheme\_module\_name returns a symbol), then place the shared object in a special directory with a special name, so that it is detected by the module loader when require is used. The special directory is a platform-specific path that can be obtained by evaluating (build-path "compiled" "native" (system-library-subpath)); see load/use-compiled for more information. For example, if the shared object's name is "example\_ss.dll", then (require "example.ss") will be redirected to "example\_ss.dll" if the latter is placed in the sub-directory (build-path "compiled" "native" (system-library-subpath)) and if "example.ss" does not exist or has an earlier timestamp.

Note that (load-extension path) within a module does not introduce the extension's definitions into the module, because load-extension is a run-time operation. To introduce an extension's bindings into a module, make sure that the extension defines a module, put the extension in the platform-specific location as described above, and use require.

**IMPORTANT:** With PLT Scheme CGC, Scheme values are garbage collected using a conservative garbage collector, so pointers to Scheme objects can be kept in registers, stack variables, or structures allocated with scheme\_malloc. However, static variables that contain pointers to collectable memory must be registered using scheme\_register\_extension\_global (see §3 "Memory Allocation").

As an example, the following C code defines an extension that returns "hello world" when it is loaded:

```
#include "escheme.h"
Scheme_Object *scheme_initialize(Scheme_Env *env) {
   return scheme_make_utf8_string("hello world");
}
Scheme_Object *scheme_reload(Scheme_Env *env) {
   return scheme_initialize(env); /* Nothing special for reload */
}
Scheme_Object *scheme_module_name() {
   return scheme_false;
}
```

Assuming that this code is in the file "hw.c", the extension is compiled under Unix with the following two commands:

```
mzc --cgc --cc hw.c
mzc --cgc --ld hw.so hw.o
```

(Note that the --cgc, --cc, and --ld flags are each prefixed by two dashes, not one.)

The "collects/mzscheme/examples" directory in the PLT distribution contains additional examples.

#### 1.2.2 3m Extensions

To build an extension to work with PLT Scheme 3m, the CGC instructions must be extended as follows:

• Adjust code to cooperate with the garbage collector as described in §3.1 "Cooperating

with 3m". Using mzc with the --xform might convert your code to implement part of the conversion, as described in §3.1.3 "Local Pointers and mzc --xform".

- In either your source in the in compiler command line, #define MZ\_PRECISE\_GC before including "escheme.h". When using mzc with the --cc and --3m flags, MZ\_PRECISE\_GC is automatically defined.
- Link with "mzdyn3m.o" (Unix, Mac OS X) or "mzdyn3m.obj" (Windows) to create a shared object. When using mzc, use the --ld and --3m flags to link to these libraries.

For a relatively simple extension "hw.c", the extension is compiled under Unix for 3m with the following three commands:

```
mzc --xform --cc hw.c
mzc --3m --cc hw.3m.c
mzc --3m --ld hw.so hw.o
```

Some examples in "collects/mzscheme/examples" work with MzScheme3m in this way. A few examples are manually instrumented, in which case the --xform step should be skipped.

# 1.3 Embedding MzScheme into a Program

Like creating extensions, the embedding process for PLT Scheme CGC or PLT Scheme 3m is essentially the same, but the process for PLT Scheme 3m is most easily understood as a variant of the process for PLT Scheme CGC.

#### 1.3.1 CGC Embedding

To embed PLT Scheme CGC in a program, follow these steps:

• Locate or build the PLT Scheme CGC libraries. Since the standard distribution provides 3m libraries, only, you will most likely have to build from source.

Under Unix, the libraries are "libmzscheme.a" and "libmzgc.a" (or "libmzscheme.so" and "libmzgc.so" for a dynamic-library build, with "libmzscheme.la" and "libmzgc.la" files for use with libtool). Building from source and installing places the libraries into the installation's "lib" directory. Be sure to build the CGC variant, since the default is 3m.

Under Windows, stub libraries for use with Microsoft tools are "libmzschx.lib" and "libmzgcx.lib" (where x represents the version number) are in a compiler-

specific directory in "plt\lib". These libraries identify the bindings that are provided by "libmzschx.dll" and "libmzgcx.dll" — which are typically installed in "plt\lib". When linking with Cygwin, link to "libmzschx.dll" and "libmzgcx.dll" directly. At run time, either "libmzschx.dll" and "libmzgcx.dll" must be moved to a location in the standard DLL search path, or your embedding application must "delayload" link the DLLs and explicitly load them before use. ("MzScheme.exe" and "MrEd.exe" use the latter strategy.)

Under Mac OS X, dynamic libraries are provided by the "PLT\_MzScheme" framework, which is typically installed in "lib" sub-directory of the installation. Supply -framework PLT\_MzScheme to gcc when linking, along with -F and a path to the "lib" directory. Beware that CGC and 3m libraries are installed as different versions within a single framework, and installation marks one version or the other as the default (by setting symbolic links); install only CGC to simplify accessing the CGC version within the framework. At run time, either "PLT\_MzScheme.framework" must be moved to a location in the standard framework search path, or your embedding executable must provide a specific path to the framework (possibly an executable-relative path using the Mach-O @executable\_path prefix).

• For each C/C++ file that uses MzScheme library functions, #include the file "scheme.h".

The C preprocessor symbol SCHEME\_DIRECT\_EMBEDDED is defined as 1 when "scheme.h" is #included, or as 0 when "escheme.h" is #included.

The "scheme.h" file is distributed with the PLT software in the installation's "include" directory. Building and installing from source also places this file in the installation's "include" directory.

- Start your main program through the scheme\_main\_setup trampoline, and put all uses of MzScheme functions inside the function passed to scheme\_main\_setup. The scheme\_main\_setup function registers the current C stack location with the memory manager. It also creates the initial namespace Scheme\_Env\* by calling scheme\_basic\_env and passing the result to the function provided to scheme\_main\_setup.
- Configure the namespace by adding module declarations. The initial namespace contains declarations only for a few primitive modules, such as '#%kernel, and no bindings are imported into the top-level environment.

To embed a module like scheme/base (along with all its dependencies), use mzc --c-mods, which generates a C file that contains modules in bytecode form as encapsulated in a static array. The generated C file defines a declare\_modules function that takes a Scheme\_Env\*, installs the modules into the environment, and adjusts the module name resolver to access the embedded declarations.

Alternately, use scheme\_set\_collects\_path and scheme\_init\_collection\_paths to configure and install a path for finding modules at run time.

- Access Scheme through scheme\_dynamic\_require, scheme\_load, scheme\_eval, and/or other functions described in this manual.
- Compile the program and link it with the MzScheme libraries.

With PLT Scheme CGC, Scheme values are garbage collected using a conservative garbage collector, so pointers to Scheme objects can be kept in registers, stack variables, or structures allocated with scheme\_malloc. In an embedding application on some platforms, static variables are also automatically registered as roots for garbage collection (but see notes below specific to Mac OS X and Windows).

For example, the following is a simple embedding program which evaluates all expressions provided on the command line and displays the results, then runs a read-eval-print loop. Run

```
mzc --c-mods base.c ++lib scheme/base
```

to generate "base.c", which encapsulates scheme/base and all of its transitive imports (so that they need not be found separately a run time).

```
#include "scheme.h"
#include "base.c"
static int run(Scheme_Env *e, int argc, char *argv[])
 Scheme_Object *curout;
  int i;
 mz_jmp_buf * volatile save, fresh;
  /* Declare embedded modules in "base.c": */
 declare_modules(e);
  scheme_namespace_require(scheme_intern_symbol("scheme/base"));
  curout = scheme_get_param(scheme_current_config(),
                            MZCONFIG_OUTPUT_PORT);
 for (i = 1; i < argc; i++) {
    save = scheme_current_thread->error_buf;
    scheme_current_thread->error_buf = &fresh;
    if (scheme_setjmp(scheme_error_buf)) {
      scheme_current_thread->error_buf = save;
      return -1; /* There was an error */
    } else {
      Scheme_Object *v, *a[2];
```

```
v = scheme_eval_string(argv[i], e);
    scheme_display(v, curout);
    scheme_display(scheme_make_char('\n'), curout);
    /* read-eval-print loop, uses initial Scheme_Env: */
    a[0] = scheme_intern_symbol("scheme/base");
    a[1] = scheme_intern_symbol("read-eval-print-loop");
    scheme_apply(scheme_dynamic_require(2, a), 0, NULL);
    scheme_current_thread->error_buf = save;
    }
}
return 0;

int main(int argc, char *argv[])
{
    return scheme_main_setup(1, run, argc, argv);
}
```

Under Mac OS X, or under Windows when MzScheme is compiled to a DLL using Cygwin, the garbage collector cannot find static variables automatically. In that case, scheme\_main\_setup must be called with a non-zero first argument.

Under Windows (for any other build mode), the garbage collector finds static variables in an embedding program by examining all memory pages. This strategy fails if a program contains multiple Windows threads; a page may get unmapped by a thread while the collector is examining the page, causing the collector to crash. To avoid this problem, call scheme\_main\_setup with a non-zero first argument.

When an embedding application calls scheme\_main\_setup with a non-zero first argument, it must register each of its static variables with MZ\_REGISTER\_STATIC if the variable can contain a GCable pointer. For example, if curout above is made static, then MZ\_REGISTER\_STATIC(curout) should be inserted before the call to scheme\_get\_param.

When building an embedded MzSchemeCGC to use SenoraGC (SGC) instead of the default collector, scheme\_main\_setup must be called with a non-zero first argument. See §3 "Memory Allocation" for more information.

#### 1.3.2 3m Embedding

MzScheme3m can be embedded mostly the same as MzScheme, as long as the embedding program cooperates with the precise garbage collector as described in §3.1 "Cooperating with 3m".

In either your source in the in compiler command line, #define MZ\_PRECISE\_GC before including "scheme.h". When using mzc with the --cc and --3m flags, MZ\_PRECISE\_GC is

automatically defined.

In addition, some library details are different:

- Under Unix, the library is just "libmzscheme3m.a" (or "libmzscheme3m.so" for a dynamic-library build, with "libmzscheme3m.la" for use with libtool). There is no separate library for 3m analogous to CGC's "libmzgc.a".
- Under Windows, the stub library for use with Microsoft tools is "libmzsch3mx.lib" (where x represents the version number). This library identifies the bindings that are provided by "libmzsch3mx.dll". There is no separate library for 3m analogous to CGC's "libmzgcx.lib".
- Under Mac OS X, 3m dynamic libraries are provided by the "PLT\_MzScheme" framework, just as for CGC, but as a version suffixed with "\_3m".

For MzScheme3m, an embedding application must call scheme\_main\_setup with a non-zero first argument.

The simple embedding program from the previous section can be processed by mzc --xform, then compiled and linked with MzScheme3m. Alternately, the source code can be extended to work with either CGC or 3m depending on whether MZ\_PRECISE\_GC is defined on the compiler's command line:

```
#include "scheme.h"
#include "base.c"
static int run(Scheme_Env *e, int argc, char *argv[])
 Scheme_Object *curout = NULL, *v = NULL, *a[2] = {NULL, NULL};
 Scheme_Config *config = NULL;
  int i;
 mz_jmp_buf * volatile save = NULL, fresh;
 MZ_GC_DECL_REG(8);
 MZ_GC_VAR_IN_REG(0, e);
 MZ_GC_VAR_IN_REG(1, curout);
 MZ_GC_VAR_IN_REG(2, save);
 MZ_GC_VAR_IN_REG(3, config);
 MZ_GC_VAR_IN_REG(4, v);
 MZ_GC_ARRAY_VAR_IN_REG(5, a, 2);
 MZ_GC_REG();
 declare_modules(e);
```

```
v = scheme_intern_symbol("scheme/base");
  scheme_namespace_require(v);
  config = scheme_current_config();
  curout = scheme_get_param(config, MZCONFIG_OUTPUT_PORT);
 for (i = 1; i < argc; i++) {
    save = scheme_current_thread->error_buf;
    scheme_current_thread->error_buf = &fresh;
    if (scheme_setjmp(scheme_error_buf)) {
      scheme_current_thread->error_buf = save;
     return -1; /* There was an error */
    } else {
      v = scheme_eval_string(argv[i], e);
      scheme_display(v, curout);
      v = scheme_make_char('\n');
      scheme_display(v, curout);
      /* read-eval-print loop, uses initial Scheme_Env: */
      a[0] = scheme_intern_symbol("scheme/base");
      a[1] = scheme_intern_symbol("read-eval-print-loop");
      v = scheme_dynamic_require(2, a);
      scheme_apply(v, 0, NULL);
      scheme_current_thread->error_buf = save;
 MZ_GC_UNREG();
 return 0;
int main(int argc, char *argv[])
 return scheme_main_setup(1, run, argc, argv);
```

Strictly speaking, the config and v variables above need not be registered with the garbage collector, since their values are not needed across function calls that allocate. The code is much easier to maintain, however, when all local variables are registered and when all temporary values are put into variables.

#### 1.4 MzScheme and Threads

MzScheme implements threads for Scheme programs without aid from the operating system, so that MzScheme threads are cooperative from the perspective of C code. Under Unix, stand-alone MzScheme uses a single OS-implemented thread. Under Windows and Mac OS X, stand-alone MzScheme uses a few private OS-implemented threads for background tasks, but these OS-implemented threads are never exposed by the MzScheme API.

In an embedding application, MzScheme can co-exist with additional OS-implemented threads, but the additional OS threads must not call any scheme\_function. Only the OS thread that originally calls scheme\_basic\_env can call scheme\_functions. (This restriction is stronger than saying all calls must be serialized across threads. MzScheme relies on properties of specific threads to avoid stack overflow and garbage collection.) When scheme\_basic\_env is called a second time to reset the interpreter, it can be called in an OS thread that is different from the original call to scheme\_basic\_env. Thereafter, all calls to scheme\_functions must originate from the new thread.

See §8 "Threads" for more information about threads, including the possible effects of MzScheme's thread implementation on extension and embedding C code.

# 1.5 MzScheme, Unicode, Characters, and Strings

A character in MzScheme is a Unicode code point. In C, a character value has type mzchar, which is an alias for unsigned — which is, in turn, 4 bytes for a properly compiled MzScheme. Thus, a mzchar\* string is effectively a UCS-4 string.

Only a few MzScheme functions use mzchar\*. Instead, most functions accept char\* strings. When such byte strings are to be used as a character strings, they are interpreted as UTF-8 encodings. A plain ASCII string is always acceptable in such cases, since the UTF-8 encoding of an ASCII string is itself.

See also §2.3 "Strings" and §11 "String Encodings".

## 1.6 Integers

MzScheme expects to be compiled in a mode where short is a 16-bit integer, int is a 32-bit integer, and long has the same number of bits as void\*. The mzlonglong type has 64 bits for compilers that support a 64-bit integer type, otherwise it is the same as long; thus, mzlonglong tends to match long long. The umzlonglong type is the unsigned version of mzlonglong.

# 2 Values and Types

A Scheme value is represented by a pointer-sized value. The low bit is a mark bit: a 1 in the low bit indicates an immediate integer, a 0 indicates a (word-aligned) pointer.

A pointer Scheme value references a structure that begins with a Scheme\_Object substructure, which in turn starts with a tag that has the C type Scheme\_Type. The rest of the structure, following the Scheme\_Object header, is type-dependent. PLT Scheme's C interface gives Scheme values the type Scheme\_Object\*. (The "object" here does not refer to objects in the sense of the scheme/class library.)

Examples of Scheme\_Type values include scheme\_pair\_type and scheme\_symbol\_type. Some of these are implemented as instances of Scheme\_Simple\_Object, which is defined in "scheme.h", but extension or embedding code should never access this structure directly. Instead, the code should use macros, such as SCHEME\_CAR, that provide access to the data of common Scheme types.

For most Scheme types, a constructor is provided for creating values of the type. For example, scheme\_make\_pair takes two Scheme\_Object\* values and returns the cons of the values.

The macro SCHEME\_TYPE takes a Scheme\_Object \* and returns the type of the object. This macro performs the tag-bit check, and returns scheme\_integer\_type when the value is an immediate integer; otherwise, SCHEME\_TYPE follows the pointer to get the type tag. Macros are provided to test for common Scheme types; for example, SCHEME\_PAIRP returns 1 if the value is a cons cell, 0 otherwise.

In addition to providing constructors, PLT Scheme defines six global constant Scheme values: scheme\_true, scheme\_false, scheme\_null, scheme\_eof, scheme\_void, and scheme\_undefined. Each of these has a type tag, but each is normally recognized via its constant address.

An extension or embedding application can create new a primitive data type by calling scheme\_make\_type, which returns a fresh Scheme\_Type value. To create a collectable instance of this type, allocate memory for the instance with scheme\_malloc. From PLT Scheme's perspective, the main constraint on the data format of such an instance is that the first sizeof(Scheme\_Object) bytes must correspond to a Scheme\_Object record; furthermore, the first sizeof(Scheme\_Type) bytes must contain the value returned by scheme\_make\_type. Extensions with modest needs can use scheme\_make\_cptr, instead of creating an entirely new type.

Scheme values should never be allocated on the stack, and they should never contain pointers to values on the stack. Besides the problem of restricting the value's lifetime to that of the stack frame, allocating values on the stack creates problems for continuations and threads, both of which copy into and out of the stack.

# 2.1 Standard Types

The following are the Scheme\_Type values for the standard types:

- scheme\_bool\_type the constants scheme\_true and scheme\_false are the
  only values of this type; use SCHEME\_FALSEP to recognize scheme\_false and use
  SCHEME\_TRUEP to recognize anything except scheme\_false; test for this type with
  SCHEME\_BOOLP
- scheme\_char\_type SCHEME\_CHAR\_VAL extracts the character (of type mzchar);
   test for this type with SCHEME\_CHARP
- scheme\_integer\_type fixnum integers, which are identified via the tag bit rather than following a pointer to this Scheme\_Type value; SCHEME\_INT\_VAL extracts the integer; test for this type with SCHEME\_INTP
- scheme\_double\_type flonum inexact numbers; SCHEME\_FLOAT\_VAL or SCHEME\_DBL\_VAL extracts the floating-point value; test for this type with SCHEME\_DBLP
- scheme\_float\_type single-precision flonum inexact numbers, when specifically enabled when compiling PLT Scheme; SCHEME\_FLOAT\_VAL or SCHEME\_FLT\_VAL extracts the floating-point value; test for this type with SCHEME\_FLTP
- scheme\_bignum\_type test for this type with SCHEME\_BIGNUMP
- scheme\_rational\_type test for this type with SCHEME\_RATIONALP
- scheme\_complex\_type test for this type or scheme\_complex\_izi\_type with SCHEME\_COMPLEXP
- scheme\_complex\_izi\_type complex number with an inexact zero imaginary part (so it counts as a real number); test for this type specifically with SCHEME\_COMPLEX\_IZIP
- scheme\_char\_string\_type SCHEME\_CHAR\_STR\_VAL extracts the string as a mzchar\*; the string is always nul-terminated, but may also contain embedded nul characters, and the Scheme string is modified if this string is modified; SCHEME\_CHAR\_STRLEN\_VAL extracts the string length (in characters, not counting the nul terminator); test for this type with SCHEME\_CHAR\_STRINGP
- scheme\_byte\_string\_type SCHEME\_BYTE\_STR\_VAL extracts the string as a char\*; the string is always nul-terminated, but may also contain embedded nul characters, and the Scheme string is modified if this string is modified; SCHEME\_BYTE\_STRLEN\_VAL extracts the string length (in bytes, not counting the nul terminator); test for this type with SCHEME\_BYTE\_STRINGP
- scheme\_path\_type SCHEME\_PATH\_VAL extracts the path as a char\*; the string is always nul-terminated; SCHEME\_PATH\_LEN extracts the path length (in bytes, not counting the nul terminator); test for this type with SCHEME\_PATHP

- scheme\_symbol\_type SCHEME\_SYM\_VAL extracts the symbol's string as a char\* UTF-8 encoding (do not modify this string); SCHEME\_SYM\_LEN extracts the number of bytes in the symbol name (not counting the nul terminator); test for this type with SCHEME\_SYMBOLP; 3m: see §3.1 "Cooperating with 3m" for a caution about SCHEME\_SYM\_VAL
- scheme\_keyword\_type SCHEME\_KEYWORD\_VAL extracts the keywors's string (without the leading hash colon) as a char\* UTF-8 encoding (do not modify this string); SCHEME\_KEYWORD\_LEN extracts the number of bytes in the keyword name (not counting the nul terminator); test for this type with SCHEME\_KEYWORDP; 3m: see §3.1 "Cooperating with 3m" for a caution about SCHEME\_KEYWORD\_VAL
- scheme\_box\_type SCHEME\_BOX\_VAL extracts/sets the boxed value; test for this type with SCHEME\_BOXP
- scheme\_pair\_type SCHEME\_CAR extracts/sets the "car" and SCHEME\_CDR extracts/sets the "cdr"; test for this type with SCHEME\_PAIRP
- scheme\_vector\_type SCHEME\_VEC\_SIZE extracts the length and SCHEME\_VEC\_ELS extracts the array of Scheme values (the Scheme vector is modified when this array is modified); test for this type with SCHEME\_VECTORP; 3m: see §3.1 "Cooperating with 3m" for a caution about SCHEME\_VEC\_ELS
- scheme\_structure\_type structure instances; test for this type with SCHEME\_STRUCTP
- scheme\_struct\_type\_type structure types; test for this type with SCHEME\_STRUCT\_TYPEP
- scheme\_struct\_property\_type structure type properties
- scheme\_input\_port\_type SCHEME\_INPORT\_VAL extracts/sets the user data pointer; test for just this type with SCHEME\_INPORTP, but use SCHEME\_INPUT\_PORTP to recognize all input ports (including structures with the prop:input-port property)
- scheme\_output\_port\_type SCHEME\_OUTPORT\_VAL extracts/sets the user data pointer; test for just this type with SCHEME\_OUTPORTP, but use SCHEME\_OUTPUT\_PORTP to recognize all output ports (including structures with the prop:output-port property)
- scheme\_thread\_type thread descriptors; test for this type with SCHEME\_THREADP
- scheme\_sema\_type semaphores; test for this type with SCHEME\_SEMAP
- scheme\_hash\_table\_type test for this type with SCHEME\_HASHTP
- scheme\_bucket\_table\_type test for this type with SCHEME\_BUCKTP

- scheme\_weak\_box\_type test for this type with SCHEME\_WEAKP; SCHEME\_WEAK\_PTR
  extracts the contained object, or NULL after the content is collected; do not set the
  content of a weak box
- scheme\_namespace\_type namespaces; test for this type with SCHEME\_NAMESPACEP
- scheme\_cpointer\_type #<void> pointer with a type-describing Scheme\_Object; SCHEME\_CPTR\_VAL extracts the pointer and SCHEME\_CPTR\_TYPE extracts the type tag object; test for this type with SCHEME\_CPTRP. The tag is used when printing such objects when it's a symbol, a byte string, a string, or a pair holding one of these in its car.

The following are the procedure types:

- scheme\_prim\_type a primitive procedure, possibly with data elements
- scheme\_closed\_prim\_type an old-style primitive procedure with a data pointer
- scheme\_compiled\_closure\_type a Scheme procedure
- scheme\_cont\_type a continuation
- scheme\_escaping\_cont\_type an escape continuation
- scheme\_case\_closure\_type a case-lambda procedure
- scheme\_native\_closure\_type a procedure with native code generated by the just-in-time compiler

}

The predicate SCHEME\_PROCP returns 1 for all procedure types and 0 for anything else.

The following are additional number predicates:

- SCHEME\_NUMBERP all numerical types
- SCHEME\_REALP all non-complex numerical types, plus scheme\_complex\_izi\_type
- SCHEME\_EXACT\_INTEGERP fixnums and bignums
- SCHEME\_EXACT\_REALP fixnums, bignums, and rationals
- SCHEME\_FLOATP both single-precision (when enabled) and double-precision flonums

### 2.2 Global Constants

There are six global constants:

- scheme\_null test for this value with SCHEME\_NULLP
- scheme\_eof test for this value with SCHEME\_EOFP
- scheme\_true
- scheme\_false test for this value with SCHEME\_FALSEP; test against it with SCHEME\_TRUEP
- scheme\_void test for this value with SCHEME\_VOIDP
- scheme\_undefined

# 2.3 Strings

As noted in §1.5 "MzScheme, Unicode, Characters, and Strings", a Scheme character is a Unicode code point represented by a mzchar value, and character strings are mzchar arrays. PLT Scheme also supplies byte strings, which are char arrays.

For a character string s, SCHEME\_CHAR\_STR\_VAL(s) produces a pointer to mzchars, not chars. Convert a character string to its UTF-8 encoding as byte string with scheme\_char\_string\_to\_byte\_string. For a byte string bs, SCHEME\_BYTE\_STR\_VAL(bs) produces a pointer to chars. The function scheme\_byte\_string\_to\_char\_string decodes a byte string as UTF-8 and produces a character string. The functions scheme\_char\_string\_to\_byte\_string\_locale and scheme\_byte\_string\_to\_char\_string\_locale are similar, but they use the current locale's encoding instead of UTF-8.

For more fine-grained control over UTF-8 encoding, use the scheme\_utf8\_decode and scheme\_utf8\_encode functions, which are described in §11 "String Encodings".

## 2.4 Value Functions

Scheme\_Object\* scheme\_make\_char(mzchar ch)

Returns the character value. The *ch* value must be a legal Unicode code point (and not a surrogate, for example). The first 256 characters are represented by constant Scheme values, and others are allocated.

```
Scheme_Object* scheme_make_char_or_null(mzchar ch)
Like scheme_make_char, but the result is NULL if ch is not a legal Unicode code point.
Scheme_Object* scheme_make_character(mzchar ch)
Returns the character value. This is a macro that directly accesses the array of constant
characters when ch is less than 256.
Scheme_Object* scheme_make_ascii_character(mzchar ch)
Returns the character value, assuming that ch is less than 256. (This is a macro.)
Scheme_Object* scheme_make_integer(long i)
Returns the integer value; i must fit in a fixnum. (This is a macro.)
Scheme_Object* scheme_make_integer_value(long i)
Returns the integer value. If i does not fit in a fixnum, a bignum is returned.
 Scheme_Object*
scheme_make_integer_value_from_unsigned(unsigned long i)
Like scheme_make_integer_value, but for unsigned integers.
 Scheme_Object*
scheme_make_integer_value_from_long_long(mzlonglong i)
Like scheme_make_integer_value, but for mzlonglong values (see §1.6 "Integers").
 Scheme_Object*
scheme\_make\_integer\_value\_from\_unsigned\_long\_long(umzlonglong\ i)
Like scheme_make_integer_value_from_long_long, but for unsigned integers.
 Scheme_Object*
scheme_make_integer_value_from_long_halves(unsigned long hi,
```

unsigned long lo)

Creates an integer given the high and low longs of a signed integer. Note that on 64-bit platforms where long long is the same as long, the resulting integer has 128 bits. (See also  $\S1.6$  "Integers".)

```
\label{lem:scheme_object} $$ scheme_make_integer_value\_from\_unsigned\_long\_halves(unsigned\_long\_hi, unsigned\_long\_lo) $$ unsigned\_long\_lo)$
```

Creates an integer given the high and low longs of an unsigned integer. Note that on 64-bit platforms where long long is the same as long, the resulting integer has 128 bits.

Extracts the integer value. Unlike the SCHEME\_INT\_VAL macro, this procedure will extract an integer that fits in a long from a Scheme bignum. If o fits in a long, the extracted integer is placed in \*i and 1 is returned; otherwise, 0 is returned and \*i is unmodified.

Like scheme\_get\_int\_val, but for unsigned integers.

Like scheme\_get\_int\_val, but for mzlonglong values (see §1.6 "Integers").

Like scheme\_get\_int\_val, but for unsigned mzlonglong values (see §1.6 "Integers").

```
Scheme_Object* scheme_make_double(double d)
```

Creates a new floating-point value.

```
Scheme_Object* scheme_make_float(float d)
```

Creates a new single-precision floating-point value. The procedure is available only when

PLT Scheme is compiled with single-precision numbers enabled.

```
double scheme_real_to_double(Scheme_Object* o)
```

Converts a Scheme real number to a double-precision floating-point value.

```
Scheme_Object* scheme_make_pair(Scheme_Object* carv, Scheme_Object* cdrv)
```

Makes a \scmi{cons} pair.

```
Scheme_Object* scheme_make_byte_string(char* bytes)
```

Makes a Scheme byte string from a nul-terminated C string. The bytes string is copied.

```
Scheme_Object*
scheme_make_byte_string_without_copying(char* bytes)
```

Like scheme\_make\_byte\_string, but the string is not copied.

Makes a byte string value with size *len*. A copy of *bytes* is made if *copy* is not 0. The string *bytes* should contain *len* bytes; *bytes* can contain the nul byte at any position, and need not be nul-terminated if *copy* is non-zero. However, if *len* is negative, then the nul-terminated length of *bytes* is used for the length, and if *copy* is zero, then *bytes* must be nul-terminated.

```
Scheme_Object* scheme_make_sized_offset_byte_string(char* bytes, long d, long len, int copy)
```

Like scheme\_make\_sized\_byte\_string, except the *len* characters start from position *d* in *bytes*. If *d* is non-zero, then *copy* must be non-zero.

Allocates a new Scheme byte string.

```
Scheme_Object* scheme_append_byte_string(Scheme_Object* a, Scheme_Object* b)
```

Creates a new byte string by appending the two given byte strings.

```
Scheme_Object* scheme_make_locale_string(char* bytes)
```

Makes a Scheme string from a nul-terminated byte string that is a locale-specific encoding of a character string; a new string is allocated during decoding. The "locale in the name of this function thus refers to *bytes*, and not the resulting string (which is internally stored as UCS-4).

```
Scheme_Object* scheme_make_utf8_string(char* bytes)
```

Makes a Scheme string from a nul-terminated byte string that is a UTF-8 encoding. A new string is allocated during decoding. The "utf8" in the name of this function thus refers to *bytes*, and not the resulting string (which is internally stored as UCS-4).

Makes a string value, based on *len* UTF-8-encoding bytes (so the resulting string is *len* characters or less). The string *bytes* should contain at least *len* bytes; *bytes* can contain the nul byte at any position, and need not be null-terminated. However, if *len* is negative, then the nul-terminated length of *bytes* is used for the length.

```
Scheme_Object* scheme_make_sized_offset_utf8_string(char* bytes, long d, long len)
```

Like  $scheme_make_sized_char_string$ , except the len characters start from position d in bytes.

```
Scheme_Object* scheme_make_char_string(mzchar* chars)
```

Makes a Scheme string from a nul-terminated UCS-4 string. The *chars* string is copied.

```
Scheme_Object*
scheme_make_char_string_without_copying(mzchar* chars)
```

Like scheme\_make\_char\_string, but the string is not copied.

Makes a string value with size *len*. A copy of *chars* is made if *copy* is not 0. The string *chars* should contain *len* characters; *chars* can contain the nul character at any position, and need not be nul-terminated if *copy* is non-zero. However, if *len* is negative, then the nulterminated length of *chars* is used for the length, and if *copy* is zero, then the *chars* must be nul-terminated.

```
Scheme_Object*
scheme_make_sized_offset_char_string(mzchar* chars,
long d,
long len,
int copy)
```

Like scheme\_make\_sized\_char\_string, except the *len* characters start from position *d* in *chars*. If *d* is non-zero, then *copy* must be non-zero.

```
Scheme_Object* scheme_alloc_char_string(int size, mzchar fill)
```

Allocates a new Scheme string.

```
Scheme_Object* scheme_append_char_string(Scheme_Object* a,
Scheme_Object* b)
```

Creates a new string by appending the two given strings.

```
Scheme_Object*
scheme_char_string_to_byte_string(Scheme_Object* s)
```

Converts a Scheme character string into a Scheme byte string via UTF-8.

```
Scheme_Object*
scheme_byte_string_to_char_string(Scheme_Object* s)
```

Converts a Scheme byte string into a Scheme character string via UTF-8.

```
Scheme_Object*
scheme_char_string_to_byte_string_locale(Scheme_Object* s)
```

Converts a Scheme character string into a Scheme byte string via the locale's encoding.

```
Scheme_Object*
scheme_byte_string_to_char_string_locale(Scheme_Object* s)
```

Converts a Scheme byte string into a Scheme character string via the locale's encoding.

```
Scheme_Object* scheme_intern_symbol(char* name)
```

Finds (or creates) the symbol matching the given nul-terminated, ASCII string (not UTF-8). The case of *name* is (non-destructively) normalized before interning if scheme\_case\_sensitive is 0.

Creates or finds a symbol given the symbol's length in UTF-8-encoding bytes. The the case of *name* is not normalized.

Like scheme\_intern\_exact\_symbol, but given a character array instead of a UTF-8-encoding byte array.

```
Scheme_Object* scheme_make_symbol(char* name)
```

Creates an uninterned symbol from a nul-terminated, UTF-8-encoding string. The case is not normalized.

Creates an uninterned symbol given the symbol's length in UTF-8-encoded bytes.

Creates or finds a keyword given the keywords length in UTF-8-encoding bytes. The the case of *name* is not normalized, and it should not include the leading hash and colon of the keyword's printed form.

```
Scheme_Object* scheme_intern_exact_char_keyword(mzchar* name, int. len)
```

Like scheme\_intern\_exact\_keyword, but given a character array instead of a UTF-8-encoding byte array.

Allocates a new vector.

```
Scheme_Object* scheme_box(Scheme_Object* v)
```

Creates a new box containing the value v.

```
Scheme_Object* scheme_make_weak_box(Scheme_Object* v)
```

Creates a new weak box containing the value v.

```
Scheme_Type scheme_make_type(char* name)
```

Creates a new type (not a Scheme value).

Creates a C-pointer object that encapsulates *ptr* and uses *typetag* to identify the type of the pointer. The SCHEME\_CPTRP macro recognizes objects created by scheme\_make\_cptr. The SCHEME\_CPTR\_VAL macro extracts the original *ptr* from the Scheme object, and SCHEME\_CPTR\_TYPE extracts the type tag. The SCHEME\_CPTR\_OFFSETVAL macro returns 0 for the result Scheme object.

Creates a C-pointer object that encapsulates both *ptr* and *offset*. The SCHEME\_CPTR\_OFFSETVAL macro returns *offset* for the result Scheme object (and the macro be used to change the offset, since it also works on objects with no offset).

Installs a printer to be used for printing (or writing or displaying) values that have the type tag *type*.

The type of *printer* is defined as follows:\cppIndex{scheme\_Type\_Printer}

Such a printer must print a representation of the value using scheme\_print\_bytes and scheme\_print\_string. The first argument to the printer, v, is the value to be printed. The second argument indicates whether v is printed via write or display. The last argument is to be passed on to scheme\_print\_bytes or scheme\_print\_string to identify the printing context.

Writes the content of str — starting from *offset* and running len bytes — into a printing context determined by pp. This function is for use by a printer that is installed with  $scheme\_set\_type\_printer$ .

Writes the content of *str* — starting from *offset* and running *len* characters — into a printing context determined by *pp*. This function is for use by a printer that is installed with scheme\_set\_type\_printer.

Installs an equality predicate and associated hash functions for values that have the type tag *type*. The *equalp* predicate is only applied to values that both have tag *type*.

The type of equalp, hash1, and hash2 are defined as follows:

The two hash functions are use to generate primary and secondary keys for double hashing in an equal?-based hash table. The result of the primary-key function should depend on both *obj* and *base*.

The cycle\_data argument in each case allows checking and hashing on cyclic values. It is intended for use in recursive checking or hashing via scheme\_recur\_equal, scheme\_recur\_equal\_hash\_key, and scheme\_recur\_equal\_hash\_key. That is, do not call plain scheme\_equal, scheme\_equal\_hash\_key, or scheme\_equal\_hash\_key for recursive checking or hashing on sub-elements of the given value(s).

# 3 Memory Allocation

PLT Scheme uses both malloc and allocation functions provided by a garbage collector. Embedding/extension C/C++ code may use either allocation method, keeping in mind that pointers to garbage-collectable blocks in malloced memory are invisible (i.e., such pointers will not prevent the block from being garbage-collected).

PLT Scheme CGC uses a conservative garbage collector. This garbage collector normally only recognizes pointers to the beginning of allocated objects. Thus, a pointer into the middle of a GC-allocated string will normally not keep the string from being collected. The exception to this rule is that pointers saved on the stack or in registers may point to the middle of a collectable object. Thus, it is safe to loop over an array by incrementing a local pointer variable.

PLT Scheme 3m uses a precise garbage collector that moves objects during collection, in which case the C code must be instrumented to expose local pointer bindings to the collector, and to provide tracing procedures for (tagged) records containing pointers. This instrumentation is described further in §3.1 "Cooperating with 3m".

The basic collector allocation functions are:

- scheme\_malloc Allocates collectable memory that may contain pointers to collectable objects; for 3m, the memory must be an array of pointers (though not necessarily to collectable objects). The newly allocated memory is initially zeroed.
- scheme\_malloc\_atomic Allocates collectable memory that does not contain
  pointers to collectable objects. If the memory does contain pointers, they are invisible
  to the collector and will not prevent an object from being collected. Newly allocated
  atomic memory is not necessary zeroed.
  - Atomic memory is used for strings or other blocks of memory which do not contain pointers. Atomic memory can also be used to store intentionally-hidden pointers.
- scheme\_malloc\_tagged Allocates collectable memory that contains a mixture of pointers and atomic data. With the conservative collector, this function is the same as scheme\_malloc, but under 3m, the type tag stored at the start of the block is used to determine the size and shape of the object for future garbage collection (as described in §3.1 "Cooperating with 3m").
- scheme\_malloc\_allow\_interior Allocates a large array of pointers such that references are allowed into the middle of the block under 3m, and such pointers prevent the block from being collected. This procedure is the same as scheme\_malloc with the conservative collector, but in the that case, having *only* a pointer into the interior will not prevent the array from being collected.
- scheme\_malloc\_atomic\_allow\_interior Like scheme\_malloc\_allow\_interior for memory that does not contain pointers.

• scheme\_malloc\_uncollectable — Allocates uncollectable memory that may contain pointers to collectable objects. There is no way to free the memory. The newly allocated memory is initially zeroed. This function is not available in 3m.

If a PLT Scheme extension stores Scheme pointers in a global or static variable, then that variable must be registered with scheme\_register\_extension\_global; this makes the pointer visible to the garbage collector. Registered variables need not contain a collectable pointer at all times (even with 3m, but the variable must contain some pointer, possibly uncollectable, at all times).

With conservative collection, no registration is needed for the global or static variables of an embedding program, unless it calls <code>scheme\_main\_setup</code> or <code>scheme\_set\_stack\_base</code> with a non-zero first or second (respectively) argument. In that case, global and static variables containing collectable pointers must be registered with <code>scheme\_register\_static</code>. The <code>MZ\_REGISTER\_STATIC</code> macro takes any variable name and registers it with <code>scheme\_register\_static</code>. The <code>scheme\_register\_static</code> function can be safely called even when it's not needed, but it must not be called multiple times for a single memory address.

Collectable memory can be temporarily locked from collection by using the reference-counting function scheme\_dont\_gc\_ptr. Under 3m, such locking does not prevent the object from being moved.

Garbage collection can occur during any call into Scheme or its allocator, on anytime that Scheme has control, except during functions that are documented otherwise. The predicate and accessor macros listed in §2.1 "Standard Types" never trigger a collection.

## 3.1 Cooperating with 3m

To allow 3m's precise collector to detect and update pointers during garbage collection, all pointer values must be registered with the collector, at least during the times that a collection may occur. The content of a word registered as a pointer must contain either NULL, a pointer to the start of a collectable object, a pointer into an object allocated by scheme\_malloc\_allow\_interior, a pointer to an object currently allocated by another memory manager (and therefore not into a block that is currently managed by the collector), or a pointer to an odd-numbered address (e.g., a Scheme fixnum).

Pointers are registered in three different ways:

- Pointers in static variables should be registered with scheme\_register\_static or MZ\_REGISTER\_STATIC.
- Pointers in allocated memory are registered automatically when they are in an array allocated with scheme\_malloc, etc. When a pointer resides in an object allocated

with scheme\_malloc\_tagged, etc.~the tag at the start of the object identifiers the object's size and shape. Handling of tags is described in §3.1.1 "Tagged Objects".

• Local pointers (i.e., pointers on the stack or in registers) must be registered through the MZ\_GC\_DECL\_REG, etc. macros that are described in §3.1.2 "Local Pointers".

A pointer must never refer to the interior of an allocated object (when a garbage collection is possible), unless the object was allocated with scheme\_malloc\_allow\_interior. For this reason, pointer arithmetic must usually be avoided, unless the variable holding the generated pointer is NULLed before a collection.

**IMPORTANT:** The SCHEME\_SYM\_VAL, SCHEME\_KEYWORD\_VAL, SCHEME\_VEC\_ELS, and SCHEME\_PRIM\_CLOSURE\_ELS macros produce pointers into the middle of their respective objects, so the results of these macros must not be held during the time that a collection can occur. Incorrectly retaining such a pointer can lead to a crash.

#### 3.1.1 Tagged Objects

As explained in §2 "Values and Types", the scheme\_make\_type function can be used to obtain a new tag for a new type of object. These new types are in relatively short supply for 3m; the maximum tag is 255, and Scheme itself uses nearly 200.

After allocating a new tag in 3m (and before creating instances of the tag), a *size procedure*, a *mark procedure*, and a *fixup procedure* must be installed for the tag using GC\_register\_traversers.

A size procedure simply takes a pointer to an object with the tag and returns its size in words (not bytes). The gcBYTES\_TO\_WORDS macro converts a byte count to a word count.

A mark procedure is used to trace references among objects without moving any objects. The procedure takes a pointer to an object, and it should apply the gcMARK macro to every pointer within the object. The mark procedure should return the same result as the size procedure.

A fixup procedure is used to update references to objects after or while they are moved. The procedure takes a pointer to an object, and it should apply the gcFIXUP macro to every pointer within the object; the expansion of this macro takes the address of its argument. The fixup procedure should return the same result as the size procedure.

Depending on the collector's implementation, the mark or fixup procedure might not be used. For example, the collector may only use the mark procedure and not actually move the object. Or it may use the fixup procedure to mark and move objects at the same time. To dereference an object pointer during a fixup procedure, use GC\_fixup\_self to convert the address passed to the procedure to refer to the potentially moved object, and use GC\_resolve to convert an address that is not yet fixed up to determine the object's current location.

When allocating a tagged object in 3m, the tag must be installed immediately after the object is allocated—or, at least, before the next possible collection.

#### 3.1.2 Local Pointers

The 3m collector needs to know the address of every local or temporary pointer within a function call at any point when a collection can be triggered. Beware that nested function calls can hide temporary pointers; for example, in

the result from one scheme\_make\_pair call is on the stack or in a register during the other call to scheme\_make\_pair; this pointer must be exposed to the garbage collection and made subject to update. Simply changing the code to

does not expose all pointers, since tmp must be evaluated before the second call to scheme\_make\_pair. In general, the above code must be converted to the form

```
tmp1 = scheme_make_pair(scheme_true, scheme_false);
tmp2 = scheme_make_pair(scheme_true, scheme_false);
scheme_make_pair(tmp1, tmp2);
```

and this is converted form must be instrumented to register tmp1 and tmp2. The final result might be

```
Scheme_Object *tmp1 = NULL, *tmp2 = NULL, *result;
MZ_GC_DECL_REG(2);

MZ_GC_VAR_IN_REG(0, tmp1);
MZ_GC_VAR_IN_REG(1, tmp2);
MZ_GC_REG();

tmp1 = scheme_make_pair(scheme_true, scheme_false);
tmp2 = scheme_make_pair(scheme_true, scheme_false);
result = scheme_make_pair(tmp1, tmp2);

MZ_GC_UNREG();
return result;
```

Notice that result is not registered above. The MZ\_GC\_UNREG macro cannot trigger a garbage collection, so the result variable is never live during a potential collection. Note also that tmp1 and tmp2 are initialized with NULL, so that they always contain a pointer whenever a collection is possible.

The MZ\_GC\_DECL\_REG macro expands to a local-variable declaration to hold information for the garbage collector. The argument is the number of slots to provide for registration. Registering a simple pointer requires a single slot, whereas registering an array of pointers requires three slots. For example, to register a pointer tmp and an array of 10 char\*s:

```
{
   Scheme_Object *tmp1 = NULL;
   char *a[10];
   int i;
   MZ_GC_DECL_REG(4);

   MZ_GC_ARRAY_VAR_IN_REG(0, a, 10);
   MZ_GC_VAR_IN_REG(3, tmp1);
   /* Clear a before a potential GC: */
   for (i = 0; i < 10; i++) a[i] = NULL;
   ...
   f(a);
  ...
}</pre>
```

The MZ\_GC\_ARRAY\_VAR\_IN\_REG macro registers a local array given a starting slot, the array variable, and an array size. The MZ\_GC\_VAR\_IN\_REG takes a slot and simple pointer variable. A local variable or array must not be registered multiple times.

In the above example, the first argument to MZ\_GC\_VAR\_IN\_REG is 3 because the information for a uses the first three slots. Even if a is not used after the call to f, a must be registered with the collector during the entire call to f, because f presumably uses a until it returns.

The name used for a variable need not be immediate. Structure members can be supplied as well:

```
{
  struct { void *s; int v; void *t; } x = {NULL, 0, NULL};
  MZ_GC_DECL_REG(2);

MZ_GC_VAR_IN_REG(0, x.s);
  MZ_GC_VAR_IN_REG(0, x.t);
  ...
}
```

In general, the only constraint on the second argument to MZ\_GC\_VAR\_IN\_REG or MZ\_GC\_ARRAY\_VAR\_IN\_REG is that & must produce the relevant address, and that address

must be on the stack.

Pointer information is not actually registered with the collector until the MZ\_GC\_REG macro is used. The MZ\_GC\_UNREG macro de-registers the information. Each call to MZ\_GC\_REG must be balanced by one call to MZ\_GC\_UNREG.

Pointer information need not be initialized with MZ\_GC\_VAR\_IN\_REG and MZ\_GC\_ARRAY\_VAR\_IN\_REG before calling MZ\_GC\_REG, and the set of registered pointers can change at any time—as long as all relevent pointers are registered when a collection might occur. The following example recycles slots and completely de-registers information when no pointers are relevant. The example also illustrates how MZ\_GC\_UNREG is not needed when control escapes from the function, such as when scheme\_signal\_error escapes.

```
Scheme_Object *tmp1 = NULL, *tmp2 = NULL;
mzchar *a, *b;
MZ_GC_DECL_REG(2);
MZ_GC_VAR_IN_REG(0, tmp1);
MZ_GC_VAR_IN_REG(1, tmp2);
tmp1 = scheme_make_utf8_string("foo");
MZ_GC_REG();
tmp2 = scheme_make_utf8_string("bar");
tmp1 = scheme_append_char_string(tmp1, tmp2);
if (SCHEME_FALSEP(tmp1))
  scheme_signal_error("shouldn't happen!");
a = SCHEME_CHAR_VAL(tmp1);
MZ_GC_VAR_IN_REG(0, a);
tmp2 = scheme_make_pair(scheme_read_bignum(a, 0, 10), tmp2);
MZ_GC_UNREG();
if (SCHEME_INTP(tmp2)) {
  return 0;
MZ_GC_REG();
tmp1 = scheme_make_pair(scheme_read_bignum(a, 0, 8), tmp2);
MZ_GC_UNREG();
return tmp1;
```

}

A MZ\_GC\_DECL\_REG can be used in a nested block to hold declarations for the block's variables. In that case, the nested MZ\_GC\_DECL\_REG must have its own MZ\_GC\_REG and MZ\_GC\_UNREG calls.

```
{
 Scheme_Object *accum = NULL;
 MZ_GC_DECL_REG(1);
 MZ_GC_VAR_IN_REG(0, accum);
 MZ_GC_REG();
  accum = scheme_make_pair(scheme_true, scheme_null);
    Scheme_Object *tmp = NULL;
    MZ_GC_DECL_REG(1);
    MZ_GC_VAR_IN_REG(0, tmp);
    MZ_GC_REG();
    tmp = scheme_make_pair(scheme_true, scheme_false);
    accum = scheme_make_pair(tmp, accum);
    MZ_GC_UNREG();
 accum = scheme_make_pair(scheme_true, accum);
 MZ_GC_UNREG();
 return accum;
```

Variables declared in a local block can also be registered together with variables from an enclosing block, but the local-block variable must be unregistered before it goes out of scope. The MZ\_GC\_NO\_VAR\_IN\_REG macro can be used to unregister a variable or to initialize a slot as having no variable.

```
{
    Scheme_Object *accum = NULL;
    MZ_GC_DECL_REG(2);
    MZ_GC_VAR_IN_REG(0, accum);
    MZ_GC_NO_VAR_IN_REG(1);
    MZ_GC_REG();

accum = scheme_make_pair(scheme_true, scheme_null);
    {
        Scheme_Object *tmp = NULL;
        MZ_GC_VAR_IN_REG(1, tmp);
    }
}
```

```
tmp = scheme_make_pair(scheme_true, scheme_false);
    accum = scheme_make_pair(tmp, accum);

MZ_GC_NO_VAR_IN_REG(1);
}
accum = scheme_make_pair(scheme_true, accum);

MZ_GC_UNREG();
return accum;
}
```

The MZ\_GC\_ macros all expand to nothing when MZ\_PRECISE\_GC is not defined, so the macros can be placed into code to be compiled for both conservative and precise collection.

The MZ\_GC\_REG and MZ\_GC\_UNREG macros must never be used in an OS thread other than Scheme's thread.

#### 3.1.3 Local Pointers and mzc --xform

When mzc is run with the --xform flag and a source C program, it produces a C program that is instrumented in the way described in the previous section (but with a slightly different set of macros). For each input file "name.c", the transformed output is "name.3m.c".

The --xform mode for mzc does not change allocation calls, nor does it generate size, mark, or fixup predocures. It merely converts the code to register local pointers.

Furthermore, the --xform mode for mzc does not handle all of C. It's ability to rearrange compound expressions is particularly limited, because --xform merely converts expression text heuristically instead of parsing C. A future version of the tool will correct such problems. For now, mzc in --xform mode attempts to provide reasonable error messages when it is unable to convert a program, but beware that it can miss cases. To an even more limited degree, --xform can work on C++ code. Inspect the output of --xform mode to ensure that your code is correctly instrumented.

Some specific limitations:

- The body of a for, while, or do loop must be surrounded with curly braces. (A conversion error is normally reported, otherwise.)
- Function calls may not appear on the right-hand side of an assignment within a declaration block. (A conversion error is normally reported if such an assignment is discovered.)
- Multiple function calls in ... ? ... : ... cannot be lifted. (A conversion error is normally reported, otherwise.)

- In an assignment, the left-hand side must be a local or static variable, not a field selection, pointer dereference, etc. (A conversion error is normally reported, otherwise.)
- The conversion assumes that all function calls use an immediate name for a function, as opposed to a compound expression as in s->f(). The function name need not be a top-level function name, but it must be bound either as an argument or local variable with the form *type id*; the syntax  $ret\_type$  (\*id)(...) is not recgoinzed, so bind the function type to a simple name with typedef, first: typedef  $ret\_type$  (\*type)(...); ... type id.
- Arrays and structs must be passed by address, only.
- GC-triggering code must not appear in system headers.
- Pointer-comparison expressions are not handled correctly when either of the compared expressions includes a function call. For example, a() == b() is not converted correctly when a and b produce pointer values.
- Passing the address of a local pointer to a function works only when the pointer variable remains live after the function call.
- A return; form can get converted to { stmt; return; };, which can break an if (...) return; else ... pattern.
- Local instances of union types are generally not supported.
- Pointer arithmetic cannot be converted away, and is instead reported as an error.

# 3.1.4 Guiding mzc --xform

The following macros can be used (with care!) to navigate --xform around code that it cannot handle:

• XFORM\_START\_SKIP and XFORM\_END\_SKIP: code between these two statements is ignored by the transform tool, except to tokenize it.

Example:

```
int foo(int c, ...) {
  int r = 0;
  XFORM_START_SKIP;
  {
    /* va plays strange tricks that confuse xform */
    va_list args;
    va_start(args, c);
    while (c--) {
        r += va_arg(args, int);
    }
}
```

```
}

XFORM_END_SKIP;
return r;
}
```

These macros can also be used at the top level, outside of any function. Since they have to be terminated by a semi-colon, however, top-level uses usually must be wrapped with #ifdef MZ\_PRECISE\_GC and #endif; a semi-colon by itself at the top level is not legal in C.

 XFORM\_HIDE\_EXPR: a macro that takes wraps an expression to disable processing of the expression.

Example:

```
int foo(int c, ...) {
  int r = 0;
  {
    /* va plays strange tricks that confuse xform */
    XFORM_CAN_IGNORE va_list args; /* See below */
    XFORM_HIDE_EXPR(va_start(args, c));
    while (c--) {
        r += XFORM_HIDE_EXPR(va_arg(args, int));
     }
   }
  return r;
}
```

- XFORM\_CAN\_IGNORE: a macro that acts like a type modifier (must appear first) to indicate that a declared variable can be treated as atomic. See above for an example.
- XFORM\_START\_SUSPEND and XFORM\_END\_SUSPEND: for use at the top level (outside of any function definition), and similar to XFORM\_START\_SKIP and XFORM\_END\_SKIP in that function and class bodies are not transformed. Type and prototype information is still collected for use by later transformations, however. These forms must be terminated by a semi-colon.
- XFORM\_START\_TRUST\_ARITH and XFORM\_END\_TRUST\_ARITH: for use at the top level (outside of any function definition) to disable warnings about pointer arithmetic. Use only when you're absolutely certain that the garbage collector cannot be pointers offset into the middle of a collectable object. These forms must be terminated by a semicolon.
- XFORM\_TRUST\_PLUS: a replacement for + that does not trigger pointer-arithmetic warnings. Use with care.
- XFORM\_TRUST\_MINUS: a replacement for that does not trigger pointer-arithmetic warnings. Use with care.

# 3.2 Memory Functions

```
void* scheme_malloc(size_t n)
```

Allocates n bytes of collectable memory, initially filled with zeros. In 3m, the allocated object is treated as an array of pointers.

```
void* scheme_malloc_atomic(size_t n)
```

Allocates *n* bytes of collectable memory containing no pointers visible to the garbage collector. The object is *not* initialized to zeros.

```
void* scheme_malloc_uncollectable(size_t n)
```

Non-3m, only. Allocates *n* bytes of uncollectable memory.

```
void* scheme_malloc_eternal(size_t n)
```

Allocates uncollectable atomic memory. This function is equivalent to malloc, except that the memory cannot be freed.

Allocates num \* size bytes of memory using scheme\_malloc.

```
void* scheme_malloc_tagged(size_t n)
```

Like scheme\_malloc, but in 3m, the type tag determines how the garbage collector traverses the object; see §3 "Memory Allocation".

```
void* scheme_malloc_allow_interior(size_t n)
```

Like scheme\_malloc, but in 3m, pointers are allowed to reference the middle of the object; see §3 "Memory Allocation".

```
char* scheme_strdup(char* str)
```

Copies the null-terminated string *str*; the copy is collectable.

```
char* scheme_strdup_eternal(char* str)
```

Copies the null-terminated string *str*; the copy will never be freed.

Attempts to allocate *size* bytes using *mallocf*. If the allocation fails, the exn:misc:out-of-memory exception is raised.

```
void** scheme_malloc_immobile_box(void* p)
```

Allocates memory that is not garbage-collected and that does not move (even with 3m), but whose first word contains a pointer to a collectable object. The box is initialized with p, but the value can be changed at any time. An immobile box must be explicitly freed using scheme\_free\_immobile\_box.

```
void scheme_free_immobile_box(void** b)
```

Frees an immobile box allocated with scheme\_malloc\_immobile\_box.

Registers an extension's global variable that can contain Scheme pointers. The address of the global is given in *ptr*, and its size in bytes in *size*. In addition to global variables, this function can be used to register any permanent memory that the collector would otherwise treat as atomic. A garbage collection can occur during the registration.

Initializes the GC stack base, creates the initial namespace by calling scheme\_basic\_env, and then calls *main* with the namespace, *argc*, and *argv*. (The *argc* and *argv* are just passed on to *main*, and are not inspected in any way.)

The Scheme\_Main type is defined as follows:

The result of main is the result of scheme\_main\_setup.

If *no\_auto\_statics* is non-zero, then static variables must be explicitly registered with the garbage collector; see §3 "Memory Allocation" for more information.

Overrides the GC's auto-determined stack base, and/or disables the GC's automatic traversal of global and static variables. If  $stack\_addr$  is NULL, the stack base determined by the GC is used. Otherwise, it should be the "deepest" memory address on the stack where a collectable pointer might be stored. This function should be called only once, and before any other scheme\_function is called. It never triggers a garbage collection.

The following example shows a typical use for setting the stack base for CGC:

```
int main(int argc, char **argv) {
   int dummy;
   scheme_set_stack_base(&dummy, 0);
   real_main(argc, argv); /* calls scheme_basic_env(), etc. */
}
```

Under 3m, the above code does not quite work, because *stack\_addr* must be the beginning or end of a local-frame registration. Worse, in CGC or 3m, if real\_main is declared static, the compiler may inline it and place variables containing collectable values deeper in the stack than dummy. To avoid these problems, use scheme\_main\_setup, instead.

Like scheme\_set\_stack\_base, except for the extra <code>stack\_end</code> argument. If <code>stack\_end</code> is non-NULL, then it corresponds to a point of C-stack growth after which Scheme should attempt to handle stack overflow. The <code>stack\_end</code> argument should not correspond to the actual stack end, since detecting stack overflow may take a few frames, and since handling stack overflow requires a few frames.

If <code>stack\_end</code> is NULL, then the stack end is computed automatically: the stack size assumed to be the limit reported by <code>getrlimit</code> under Unix and Mac OS X, or it is assumed to be 1 MB under Windows; if this size is greater than 8 MB, then 8 MB is assumed, instead; the size is decremented by 50000 bytes to cover a large margin of error; finally, the size is subtracted from (for stacks that grow down) or added to (for stacks that grow up) the stack base in <code>stack\_addr</code> or the auotmatically computed stack base. Note that the 50000-byte margin of error is assumed to cover the difference between the actual stack start and the reported stack base, in addition to the margin needed for detecting and handling stack overflow.

Like scheme\_register\_extension\_global, for use in embedding applications in situations where the collector does not automatically find static variables (i.e., when scheme\_set\_stack\_base has been called with a non-zero second argument).

The macro MZ\_REGISTER\_STATIC can be used directly on a static variable. It expands to a comment if statics need not be registered, and a call to scheme\_register\_static (with the address of the static variable) otherwise.

```
void scheme_weak_reference(void** p)
```

Registers the pointer p as a weak pointer; when no other (non-weak) pointers reference the same memory as p references, then p will be set to NULL by the garbage collector. The value in p may change, but the pointer remains weak with respect to the value of p at the time p was registered.

```
void scheme_weak_reference_indirect(void** p, void* v)
```

Like  $scheme_weak_reference$ , but \*p is cleared (regardless of its value) when there are no references to v.

Registers a callback function to be invoked when the memory p would otherwise be garbage-collected, and when no "will"-like finalizers are registered for p.

The fnl\_proc type is not actually defined, but it is equivalent to

```
typedef void (*fnl_proc)(void *p, void *data)
```

The f argument is the callback function; when it is called, it will be passed the value p and the data pointer data; data can be anything — it is only passed on to the callback function. If oldf and olddata are not NULL, then \*oldf and \*olddata are filled with the old callback information (f and data will override this old callback).

To remove a registered finalizer, pass NULL for f and data.

Note: registering a callback not only keeps *p* from collection until the callback is invoked, but it also keeps *data* reachable until the callback is invoked.

Adds a finalizer to a chain of primitive finalizers. This chain is separate from the single finalizer installed with scheme\_register\_finalizer; all finalizers in the chain are called immediately after a finalizer that is installed with scheme\_register\_finalizer.

See scheme\_register\_finalizer, above, for information about the arguments.

To remove an added finalizer, use scheme\_subtract\_finalizer.

Installs a "will"-like finalizer, similar to will-register. Scheme finalizers are called one at a time, requiring the collector to prove that a value has become inaccessible again before calling the next Scheme finalizer. Finalizers registered with scheme\_register\_finalizer or scheme\_add\_finalizer are not called until all Scheme finalizers have been exhausted.

See scheme\_register\_finalizer, above, for information about the arguments.

There is currently no facility to remove a "will"-like finalizer.

Like scheme\_add\_finalizer, but if the combination f and data is already registered as a (non-"will"-like) finalizer for p, it is not added a second time.

Like scheme\_add\_scheme\_finalizer, but if the combination of f and data is already registered as a "will"-like finalizer for p, it is not added a second time.

```
\begin{tabular}{ll} {\tt void} & & {\tt scheme\_subtract\_finalizer(void*~$p$,} \\ & & & {\tt fnl\_proc~$f$,} \\ & & & & {\tt void*~$data)} \end{tabular}
```

Removes a finalizer that was installed with scheme\_add\_finalizer.

```
void scheme_remove_all_finalization(void* p)
```

Removes all finalization ("will"-like or not) for *p*, including wills added in Scheme with will-register and finalizers used by custodians.

```
void scheme_dont_gc_ptr(void* p)
```

Keeps the collectable block p from garbage collection. Use this procedure when a reference to p is be stored somewhere inaccessible to the collector. Once the reference is no longer used from the inaccessible region, de-register the lock with  $scheme\_gc\_ptr\_ok$ . A garbage collection can occur during the registration.

This function keeps a reference count on the pointers it registers, so two calls to  $scheme\_dont\_gc\_ptr$  for the same p should be balanced with two calls to  $scheme\_gc\_ptr\_ok$ .

```
void scheme_gc_ptr_ok(void* p)
See scheme_dont_gc_ptr.

void scheme_collect_garbage()
```

Forces an immediate garbage-collection.

```
void GC_register_traversers(short tag,
Size_Proc s,
Mark_Proc m,
Fixup_Proc f,
int is_const_size,
int is_atomic)
```

3m only. Registers a size, mark, and fixup procedure for a given type tag; see §3.1.1 "Tagged Objects" for more information.

Each of the three procedures takes a pointer and returns an integer:

```
typedef int (*Size_Proc)(void *obj);
```

```
typedef int (*Mark_Proc)(void *obj);
typedef int (*Fixup_Proc)(void *obj);
```

If the result of the size procedure is a constant, then pass a non-zero value for *is\_const\_size*. If the mark and fixup procedures are no-ops, then pass a non-zero value for *is\_atomic*.

# 4 Namespaces and Modules

A Scheme namespace (a top-level environment) is represented by a value of type Scheme\_Env\* — which is also a Scheme value, castable to Scheme\_Object\*. Calling scheme\_basic\_env returns a namespace that includes all of Scheme's standard global procedures and syntax.

The scheme\_basic\_env function must be called once by an embedding program, before any other PLT Scheme function is called (except scheme\_make\_param), but scheme\_main\_setup automatically calls scheme\_basic\_env. The returned namespace is the initial current namespace for the main Scheme thread. Scheme extensions cannot call scheme\_basic\_env.

The current thread's current namespace is available from scheme\_get\_env, given the current parameterization (see §9 "Parameterizations"): scheme\_get\_env(scheme\_config).

New values can be added as globals in a namespace using scheme\_add\_global. The scheme\_lookup\_global function takes a Scheme symbol and returns the global value for that name, or NULL if the symbol is undefined.

A module's set of top-level bindings is implemented using the same machinery as a namespace. Use scheme\_primitive\_module to create a new Scheme\_Env\* that represents a primitive module. The name provided to scheme\_primitive\_module is subject to prefixing through the current-module-name-prefix parameter (which is normally set by the module name resolver when auto-loading module files). After installing variables into the module with scheme\_add\_global, etc., call scheme\_finish\_primitive\_module on the Scheme\_Env\* value to make the module declaration available. All defined variables are exported from the primitive module.

The Scheme #%variable-reference form produces a value that is opaque to Scheme code. Use SCHEME\_PTR\_VAL on the result of #%variable-reference to obtain the same kind of value as returned by scheme\_global\_bucket (i.e., a bucket containing the variable's value, or NULL if the variable is not yet defined).

```
void scheme_add_global(char* name,
Scheme_Object* val,
Scheme_Env* env)
```

Adds a value to the table of globals for the namespace *env*, where *name* is a null-terminated string. (The string's case will be normalized in the same way as for interning a symbol.)

Adds a value to the table of globals by symbol name instead of string name.

```
Scheme_Object* scheme_lookup_global(Scheme_Object* symbol, Scheme_Env* env)
```

Given a global variable name (as a symbol) in sym, returns the current value.

```
Scheme_Bucket* scheme_global_bucket(Scheme_Object* symbol, Scheme_Env* env)
```

Given a global variable name (as a symbol) in *sym*, returns the bucket where the value is stored. When the value in this bucket is NULL, then the global variable is undefined.

The Scheme\_Bucket structure is defined as:

```
typedef struct Scheme_Bucket {
   Scheme_Object so; /* so.type = scheme_variable_type */
   void *key;
   void *val;
} Scheme_Bucket;
```

Like scheme\_global\_bucket, but finds a variable in a module. The *mod* and *symbol* arguments are as for dynamic-require in Scheme. The *pos* argument should be -1 always. The *env* argument represents the namespace in which the module is declared.

Changes the value of a global variable. The *procname* argument is used to report errors (in case the global variable is constant, not yet bound, or bound as syntax). If *set\_undef* is not 1, then the global variable must already have a binding. (For example, set! cannot set unbound variables, while define can.)

Scheme\_Object\* scheme\_builtin\_value(const char\* name)

Gets the binding of a name as it would be defined in the initial namespace.

```
Scheme_Env* scheme_get_env(Scheme_Config* config)
```

Returns the current namespace for the given parameterization (see §9 "Parameterizations"). The current thread's current parameterization is available as scheme\_config.

```
Scheme_Env* scheme_primitive_module(Scheme_Object* name, Scheme_Env* for_env)
```

Prepares a new primitive module whose name is the symbol *name* (plus any prefix that is active via current-module-name-prefix). The module will be declared within the namespace *for\_env*. The result is a Scheme\_Env \* value that can be used with scheme\_add\_global, etc., but it represents a module instead of a namespace. The module is not fully declared until scheme\_finish\_primitive\_module is called, at which point all variables defined in the module become exported.

void scheme\_finish\_primitive\_module(Scheme\_Env\* env)

Finalizes a primitive module and makes it available for use within the module's namespace.

### 5 Procedures

A *primitive procedure* is a Scheme-callable procedure that is implemented in C. Primitive procedures are created in Scheme with the function scheme\_make\_prim\_w\_arity, which takes a C function pointer, the name of the primitive, and information about the number of Scheme arguments that it takes; it returns a Scheme procedure value.

The C function implementing the procedure must take two arguments: an integer that specifies the number of arguments passed to the procedure, and an array of Scheme\_Object\* arguments. The number of arguments passed to the function will be checked using the arity information. (The arity information provided to scheme\_make\_prim\_w\_arity is also used for the Scheme arity procedure.) The procedure implementation is not allowed to mutate the input array of arguments, although it may mutate the arguments themselves when appropriate (e.g., a fill in a vector argument).

The function scheme\_make\_prim\_closure\_w\_arity is similar to scheme\_make\_prim\_w\_arity, but it takes an additional count and Scheme\_Object\* array that is copied into the created procedure; the procedure is passed back to the C function when the closure is invoked. In this way, closure-like data from the C world can be associated with the primitive procedure.

The function scheme\_make\_closed\_prim\_w\_arity is similar to scheme\_make\_prim\_closure\_w\_arity, but it uses an older calling convention for passing closure data.

To work well with Scheme threads, a C function that performs substantial or unbounded work should occasionally call SCHEME\_USE\_FUEL; see §8.2 "Allowing Thread Switches" for details.

Creates a primitive procedure value, given the C function pointer *prim*. The form of *prim* is defined by:

The value *mina* should be the minimum number of arguments that must be supplied to the procedure. The value *maxa* should be the maximum number of arguments that can be supplied to the procedure, or -1 if the procedure can take arbitrarily many arguments. The *mina* and *maxa* values are used for automatically checking the argument count before the primitive is invoked, and also for the Scheme arity procedure. The *name* argument is used to

report application arity errors at run-time.

```
Scheme_Object* scheme_make_folding_prim(Scheme_Prim* prim, char* name, int mina, int maxa, short folding)
```

Like scheme\_make\_prim\_w\_arity, but if *folding* is non-zero, the compiler assumes that an application of the procedure to constant values can be folded to a constant. For example, +, zero?, and string-length are folding primitives, but display and cons are not.

```
Scheme_Object* scheme_make_prim(Scheme_Prim* prim)
```

Same as scheme\_make\_prim\_w\_arity, but the arity (0, -1) and the name "UNKNOWN" is assumed. This function is provided for backward compatibility only.

Creates a primitive procedure value that includes the c values in vals; when the C function prim is invoked, the generated primitive is passed as the last parameter. The form of prim is defined by:

The macro SCHEME\_PRIM\_CLOSURE\_ELS takes a primitive-closure object and returns an array with the same length and content as *vals*. (3m: see §3.1 "Cooperating with 3m" for a caution about SCHEME\_PRIM\_CLOSURE\_ELS.)

Creates an old-style primitive procedure value; when the C function *prim* is invoked, *data* is passed as the first parameter. The form of *prim* is defined by:

Creates a closed primitive procedure value without arity information. This function is provided for backward compatibility only.

## 6 Evaluation

A Scheme S-expression is evaluated by calling scheme\_eval. This function takes an S-expression (as a Scheme\_Object\*) and a namespace and returns the value of the expression in that namespace.

The function scheme\_apply takes a Scheme\_Object\* that is a procedure, the number of arguments to pass to the procedure, and an array of Scheme\_Object \* arguments. The return value is the result of the application. There is also a function scheme\_apply\_to\_list, which takes a procedure and a list (constructed with scheme\_make\_pair) and performs the Scheme apply operation.

The scheme\_eval function actually calls scheme\_compile followed by scheme\_eval\_compiled.

## **6.1 Top-level Evaluation Functions**

The functions scheme\_eval, scheme\_apply, etc., are *top-level evaluation functions*. Continuation invocations are confined to jumps within a top-level evaluation.

The functions \_scheme\_eval\_compiled, \_scheme\_apply, etc. (with a leading underscore) provide the same functionality without starting a new top-level evaluation; these functions should only be used within new primitive procedures. Since these functions allow full continuation hops, calls to non-top-level evaluation functions can return zero or multiple times.

Currently, escape continuations and primitive error escapes can jump out of all evaluation and application functions. For more information, see §7 "Exceptions and Escape Continuations".

#### **6.2** Tail Evaluation

All of Scheme's built-in functions and syntax support proper tail-recursion. When a new primitive procedure or syntax is added to Scheme, special care must be taken to ensure that tail recursion is handled properly. Specifically, when the final return value of a function is the result of an application, then scheme\_tail\_apply should be used instead of scheme\_apply. When scheme\_tail\_apply is called, it postpones the procedure application until control returns to the Scheme evaluation loop.

For example, consider the following implementation of a thunk-or primitive, which takes any number of thunks and performs or on the results of the thunks, evaluating only as many thunks as necessary.

 $\verb|static Scheme_Object *|$ 

```
thunk_or (int argc, Scheme_Object **argv)
{
  int i;
  Scheme_Object *v;

  if (!argc)
    return scheme_false;

  for (i = 0; i < argc - 1; i++)
    if (SCHEME_FALSEP((v = _scheme_apply(argv[i], 0, NULL))))
    return v;

  return scheme_tail_apply(argv[argc - 1], 0, NULL);
}</pre>
```

This thunk-or properly implements tail-recursion: if the final thunk is applied, then the result of thunk-or is the result of that application, so scheme\_tail\_apply is used for the final application.

# **6.3** Multiple Values

A primitive procedure can return multiple values by returning the result of calling scheme\_values. The functions scheme\_eval\_compiled\_multi, scheme\_apply\_multi, \_scheme\_eval\_compiled\_multi, and \_scheme\_apply\_multi potentially return multiple values; all other evaluation and applications procedures return a single value or raise an exception.

Multiple return values are represented by the scheme\_multiple\_values "value." This quasi-value has the type Scheme\_Object\*, but it is not a pointer or a fixnum. When the result of an evaluation or application is scheme\_multiple\_values, the number of actual values can be obtained as scheme\_multiple\_count, and the array of Scheme\_Object\* values as scheme\_multiple\_array. (Both of those identifiers are actually macros.)

A garbage collection must not occur between the return of a scheme\_multiple\_values "value" and the receipt of the values through scheme\_multiple\_count scheme\_multiple\_array. Furthermore, if scheme\_multiple\_array is to be used across a potential garbage collection, then it must be specifically received by calling scheme\_detach\_multiple\_array; otherwise, a garbage collection or further evaluation may change the content of the array. Otherwise, if any application or evaluation procedure is called, the scheme\_multiple\_count and scheme\_multiple\_array variables may be modified (but the array previously referenced by scheme\_multiple\_array is never re-used if scheme\_detatch\_multiple\_array is called).

The scheme\_multiple\_count and scheme\_multiple\_array variables only contain meaningful values when scheme\_multiple\_values is returned.

## **6.4** Evaluation Functions

```
Scheme_Object* scheme_eval(Scheme_Object* expr, Scheme_Env* env)
```

Evaluates the (uncompiled) S-expression *expr* in the namespace *env*.

Evaluates the compiled expression *obj*, which was previously returned from scheme\_compile, first linking to the namespace *env*.

```
Scheme_Object* scheme_eval_compiled_multi(Scheme_Object* obj, Scheme_Env* env)
```

Evaluates the compiled expression *obj*, possibly returning multiple values (see §6.3 "Multiple Values").

```
Scheme_Object* _scheme_eval_compiled(Scheme_Object* obj, Scheme_Env* env)
```

Non-top-level version of scheme\_eval\_compiled. (See §6.1 "Top-level Evaluation Functions".)

```
Scheme_Object* _scheme_eval_compiled_multi(Scheme_Object* obj, Scheme_Env* env)
```

Non-top-level version of scheme\_eval\_compiled\_multi. (See §6.1 "Top-level Evaluation Functions".)

```
Scheme_Env* scheme_basic_env()
```

Creates the main namespace for an embedded PLT Scheme. This procedure must be called before other Scheme library function (except scheme\_make\_param). Extensions to Scheme cannot call this function.

If it is called more than once, this function resets all threads (replacing the main thread), parameters, ports, namespaces, and finalizations.

Creates and returns a new namespace. This values can be cast to Scheme\_Env \*. It can also be installed in a parameterization using scheme\_set\_param with MZCONFIG\_ENV.

When PLT Scheme is embedded in an application, create the initial namespace with scheme\_basic\_env before calling this procedure to create new namespaces.

Applies the procedure f to the given arguments.

Applies the procedure f to the given arguments, possibly returning multiple values (see  $\S6.3$  "Multiple Values").

```
Scheme_Object* _scheme_apply(Scheme_Object* f, int c, Scheme_Object** args)
```

Non-top-level version of scheme\_apply. (See §6.1 "Top-level Evaluation Functions".)

```
Scheme_Object* _scheme_apply_multi(Scheme_Object* f, int c, Scheme_Object** args)
```

Non-top-level version of  $scheme\_apply\_multi$ . (See  $\S 6.1$  "Top-level Evaluation Functions".)

```
Scheme_Object* scheme_apply_to_list(Scheme_Object* f,
Scheme_Object* args)
```

Applies the procedure f to the list of arguments in args.

```
Scheme_Object* scheme_eval_string(char* str, Scheme_Env* env)
```

Reads a single S-expression from *str* and evaluates it in the given namespace; the expression must return a single value, otherwise an exception is raised. The *str* argument is parsed as a UTF-8-encoded string of Unicode characters (so plain ASCII is fine).

```
Scheme_Object* scheme_eval_string_multi(char* str,
Scheme_Env* env)
```

Like scheme\_eval\_string, but returns scheme\_multiple\_values when the expression returns multiple values.

```
Scheme_Object* scheme_eval_string_all(char* str,
Scheme_Env* env,
int all)
```

Like scheme\_eval\_string, but if *all* is not 0, then expressions are read and evaluated from *str* until the end of the string is reached.

Applies the procedure as a tail-call. Actually, this function just registers the given application to be invoked when control returns to the evaluation loop. (Hence, this function is only useful within a primitive procedure that is returning to its caller.)

Like scheme\_tail\_apply, but the array *args* is not copied. Use this only when *args* has infinite extent and will not be used again, or when *args* will certainly not be used again until the called procedure has returned.

Applies the procedure as a tail-call.

Compiles the S-expression *form* in the given namespace. The returned value can be used with scheme\_eval\_compiled et al. Provide a non-zero value fo *writable* if the resulting compiled object will be marshalled via write instead of evaluated.

Expands all macros in the S-expression form using the given namespace.

Returns the given values together as multiple return values. Unless n is 1, the result will always be scheme\_multiple\_values.

```
void scheme_detach_multiple_array(Scheme_Object** args)
```

Called to receive multiple-value results; see §6.3 "Multiple Values".

# 7 Exceptions and Escape Continuations

When Scheme encounters an error, it raises an exception. The default exception handler invokes the error display handler and then the error escape handler. The default error escape handler escapes via a *primitive error escape*, which is implemented by calling scheme\_longjmp(\*scheme\_current\_thread->error\_buf).

An embedding program should install a fresh buffer into scheme\_current\_thread->error\_buf and call scheme\_setjmp(\*scheme\_current\_thread->error\_buf) before any top-level entry into Scheme evaluation to catch primitive error escapes. When the new buffer goes out of scope, restore the original in scheme\_current\_thread->error\_buf. The macro scheme\_error\_buf is a shorthand for \*scheme\_current\_thread->error\_buf.

```
mz_jmp_buf * volatile save, fresh;
...
save = scheme_current_thread->error_buf;
scheme_current_thread->error_buf = &fresh;
if (scheme_setjmp(scheme_error_buf)) {
    /* There was an error */
    ...
} else {
    v = scheme_eval_string(s, env);
}
scheme_current_thread->error_buf = save;
...
```

3m: when scheme\_setjmp is used, the enclosing context must provide a local-variable registration record via MZ\_GC\_DECL\_REG. Use MZ\_GC\_DECL\_REG(0) if the context has no local variables to register. Unfortunately, when using --xform with mzc instead of MZ\_GC\_DECL\_REG, etc., you may need to declare a dummy pointer and use it after scheme\_setjmp to ensure that a local-variable registration is generated.

New primitive procedures can raise a generic exception by calling scheme\_signal\_error. The arguments for scheme\_signal\_error are roughly the same as for the standard C function printf. A specific primitive exception can be raised by calling scheme\_raise\_exn.

Full continuations are implemented in Scheme by copying the C stack and using scheme\_setjmp and scheme\_longjmp. As long a C/C++ application invokes Scheme evaluation through the top-level evaluation functions (scheme\_eval, scheme\_apply, etc., as opposed to \_scheme\_apply, \_scheme\_eval\_compiled, etc.), the code is protected against any unusual behavior from Scheme evaluations (such as returning twice from a function) because continuation invocations are confined to jumps within a single top-level evaluation. However, escape continuation jumps are still allowed; as explained in the following subsection, special care must be taken in extension that is sensitive to escapes.

# 7.1 Temporarily Catching Error Escapes

When implementing new primitive procedure, it is sometimes useful to catch and handle errors that occur in evaluating subexpressions. One way to do this is the following: save scheme\_current\_thread->error\_buf to a temporary variable, set scheme\_current\_thread->error\_buf to the address of a stack-allocated mz\_jmp\_buf, invoke scheme\_setjmp(scheme\_error\_buf), perform the function's work, and then restore scheme\_current\_thread->error\_buf before returning a value. (3m: A stack-allocated mz\_jmp\_buf instance need not be registered with the garbage collector, and a heap-allocated mz\_jmp\_buf should be allocated as atomic.)

However, beware that a prompt abort or the invocation of an escaping continuation looks like a primitive error escape. In that case, the special indicator flag scheme\_jumping\_to\_continuation is non-zero (instead of its normal zero value); this situation is only visible when implementing a new primitive procedure. When scheme\_jumping\_to\_continuation is non-zero, honor the escape request by chaining to the previously saved error buffer; otherwise, call scheme\_clear\_escape.

```
mz_jmp_buf * volatile save, fresh;
save = scheme_current_thread->error_buf;
scheme_current_thread->error_buf = &fresh;
if (scheme_setjmp(scheme_error_buf)) {
    /* There was an error or continuation invocation */
    if (scheme_jumping_to_continuation) {
        /* It was a continuation jump */
        scheme_longjmp(*save, 1);
        /* To block the jump, instead: scheme_clear_escape(); */
    } else {
        /* It was a primitive error escape */
    }
} else {
    scheme_eval_string("x", scheme_env);
}
scheme_current_thread->error_buf = save;
```

This solution works fine as long as the procedure implementation only calls top-level evaluation functions (scheme\_eval, scheme\_apply, etc., as opposed to \_scheme\_apply, \_scheme\_eval\_compiled, etc.). Otherwise, use scheme\_dynamic\_wind to protect your code against full continuation jumps in the same way that dynamic\_wind is used in Scheme.

The above solution simply traps the escape; it doesn't report the reason that the escape occurred. To catch exceptions and obtain information about the exception, the simplest route is to mix Scheme code with C-implemented thunks. The code below can be used to catch exceptions in a variety of situations. It implements the function <code>\_apply\_catch\_exceptions</code>, which catches exceptions during the application of a thunk. (This code is in "collects/mzscheme/examples/catch.c" in the distribution.)

```
static Scheme_Object *exn_catching_apply, *exn_p, *exn_message;
static void init_exn_catching_apply()
  if (!exn_catching_apply) {
   char *e =
      "(lambda (thunk) "
        "(with-handlers ([void (lambda (exn) (cons #f exn))]) "
          "(cons #t (thunk))))";
    /* make sure we have a namespace with the standard bindings: */
   Scheme_Env *env = (Scheme_Env *)scheme_make_namespace(0, NULL);
   scheme_register_extension_global(&exn_catching_apply,
                                      sizeof(Scheme_Object *));
   scheme_register_extension_global(&exn_p,
                                      sizeof(Scheme_Object *));
    scheme_register_extension_global(&exn_message,
                                      sizeof(Scheme_Object *));
    exn_catching_apply = scheme_eval_string(e, env);
    exn_p = scheme_lookup_global(scheme_intern_symbol("exn?"), env);
    exn_message
      = scheme_lookup_global(scheme_intern_symbol("exn-message"),
                             env);
}
/* This function applies a thunk, returning the Scheme value if
  there's no exception, otherwise returning NULL and setting *exn
   to the raised value (usually an exn structure). */
Scheme_Object *_apply_thunk_catch_exceptions(Scheme_Object *f,
                                              Scheme_Object **exn)
 Scheme_Object *v;
 init_exn_catching_apply();
 v = _scheme_apply(exn_catching_apply, 1, &f);
 /* v is a pair: (cons #t value) or (cons #f exn) */
  if (SCHEME_TRUEP(SCHEME_CAR(v)))
   return SCHEME_CDR(v);
 else {
   *exn = SCHEME_CDR(v);
   return NULL;
```

```
}
  Scheme_Object *extract_exn_message(Scheme_Object *v)
    init_exn_catching_apply();
    if (SCHEME_TRUEP(_scheme_apply(exn_p, 1, &v)))
      return _scheme_apply(exn_message, 1, &v);
      return NULL; /* Not an exn structure */
  }
In the following example, the above code is used to catch exceptions that occur during while
evaluating source code from a string.
  static Scheme_Object *do_eval(void *s, int noargc, Scheme_Object **noargv)
    return scheme_eval_string((char *)s, scheme_get_env(scheme_config));
  static Scheme_Object *eval_string_or_get_exn_message(char *s)
    Scheme_Object *v, *exn;
    v = scheme_make_closed_prim(do_eval, s);
    v = _apply_thunk_catch_exceptions(v, &exn);
    /* Got a value? */
    if (v)
      return v;
    v = extract_exn_message(exn);
    /* Got an exn? */
    if (v)
      return v;
    /* 'raise' was called on some arbitrary value */
    return exn;
```

# 7.2 Enabling and Disabling Breaks

When embedding PLT Scheme, asynchronous break exceptions are disabled by default. Call scheme\_set\_can\_break (which is the same as calling the Scheme function breakenabled) to enable or disable breaks. To enable or disable breaks during the dynamic extent

of another evaluation (where you would use with-break-parameterization in Scheme), use scheme\_push\_break\_enable before and scheme\_pop\_break\_enable after, instead.

# 7.3 Exception Functions

```
void scheme_signal_error(char* msg, ... ...)
```

Raises a generic primitive exception. The parameters are roughly as for printf, but with the following format directives:

- %c: a Unicode character (of type mzchar)
- %d: an integer
- %ld: a long integer
- %f: a floating-point double
- %s: a nul-terminated char string
- %5: a nul-terminated mzchar string
- %S: a Scheme symbol (a Scheme\_Object\*)
- %t: a char string with a long size (two arguments), possibly containing a non-terminating nul byte, and possibly without a nul-terminator
- %u: a mzchar string with a long size (two arguments), possibly containing a non-terminating nul character, and possibly without a nul-terminator
- %T: a Scheme string (a Scheme\_Object\*)
- \( \frac{1}{2} \) a string, truncated to 253 characters, with ellipses printed if the string is truncated
- 1/2 : a Scheme string (a Scheme\_Object\*), truncated to 253 characters, with ellipses printed if the string is truncated
- %V : a Scheme value (a Scheme\_Object\*), truncated according to the current error print width.
- %e: an errno value, to be printed as a text message.
- **%**E: a platform-specific error value, to be printed as a text message.
- %Z: a potential platform-specific error value and a char string; if the string is non-NULL, then the error value is ignored, otherwise the error value is used as for \Format{E}.

#### • %%: a percent sign

The arguments following the format string must include no more than 25 strings and Scheme values, 25 integers, and 25 floating-point numbers. (This restriction simplifies the implementation with precise garbage collection.)

```
void scheme_raise_exn(int exnid,
... ...)
```

Raises a specific primitive exception. The *exnid* argument specifies the exception to be raised. If an instance of that exception has *n* fields, then the next *n*-2 arguments are values for those fields (skipping the message and debug-info fields). The remaining arguments start with an error string and proceed roughly as for printf; see scheme\_signal\_error above for more details.

Exception ids are #defined using the same names as in Scheme, but prefixed with "MZ", all letters are capitalized, and all ":'s', "-"s, and "/"s are replaced with underscores. For example, MZEXN\_FAIL\_FILESYSTEM is the exception id for a filesystem exception.

```
void scheme_warning(char* msg, ... ...)
```

Signals a warning. The parameters are roughly as for printf; see scheme\_signal\_error above for more details.

This function is automatically invoked when the wrong number of arguments are given to a primitive procedure. It signals that the wrong number of parameters was received and escapes (like scheme\_signal\_error). The *name* argument is the name of the procedure that was given the wrong number of arguments; *minc* is the minimum number of expected arguments; *maxc* is the maximum number of expected arguments, or -1 if there is no maximum; *argc* and *argv* contain all of the received arguments.

Signals that an argument of the wrong type was received, and escapes (like scheme\_signal\_error). The *name* argument is the name of the procedure that was given the wrong type of argument; *expected* is the name of the expected type; *which* is the offending argument in the *argv* array; *argc* and *argv* contain all of the received arguments. If the original *argc* and *argv* are not available, provide -1 for *which* and a pointer to the bad value in *argv*; *argc* is ignored in this case.

Signals that the wrong number of values were returned to a multiple-values context. The *expected* argument indicates how many values were expected, *got* indicates the number received, and *argv* are the received values. The *detail* string can be NULL or it can contain a printf-style string (with additional arguments) to describe the context of the error; see scheme\_signal\_error above for more details about the printf-style string.

```
void scheme_unbound_global(char* name)
```

Signals an unbound-variable error, where *name* is the name of the variable.

Converts a Scheme value into a string for the purposes of reporting an error message. The *count* argument specifies how many Scheme values total will appear in the error message (so the string for this value can be scaled appropriately). If *len* is not NULL, it is filled with the length of the returned string.

Converts an array of Scheme values into a byte string, skipping the array element indicated by *which*. This function is used to specify the "other" arguments to a function when one argument is bad (thus giving the user more information about the state of the program when the error occurred). If *len* is not NULL, it is filled with the length of the returned string.

Checks the *which*th argument in *argv* to make sure it is a procedure that can take *a* arguments. If there is an error, the *where*, *which*, *argc*, and *argv* arguments are passed on to scheme\_wrong\_type. As in scheme\_wrong\_type, *which* can be -1, in which case \**argv* is checked.

Evaluates calls the function *action* to get a value for the scheme\_dynamic\_wind call. The Pre\_Post\_Proc and Action\_Proc types are not actually defined; instead the types are inlined as if they were defined as follows:

```
typedef void (*Pre_Post_Proc)(void *data);
typedef Scheme_Object* (*Action_Proc)(void *data);
```

The functions *pre* and *post* are invoked when jumping into and out of *action*, respectively.

The function *jmp\_handler* is called when an error is signaled (or an escaping continuation is invoked) during the call to *action*; if *jmp\_handler* returns NULL, then the error is passed on to the next error handler, otherwise the return value is used as the return value for the scheme\_dynamic\_wind call.

The pointer *data* can be anything; it is passed along in calls to *action*, *pre*, *post*, and *jmp\_handler*.

```
void scheme_clear_escape()
```

Clears the "jumping to escape continuation" flag associated with a thread. Call this function when blocking escape continuation hops (see the first example in §7.1 "Temporarily Catching Error Escapes").

```
void scheme_set_can_break(int on)
```

Enables or disables breaks in the same way as calling break-enabled.

Use this function with scheme\_pop\_break\_enable to enable or disable breaks in the same way as with-break-parameterization; this function writes to *cframe* to initialize it, and scheme\_pop\_break\_enable reads from *cframe*. If *pre\_check* is non-zero and breaks are currently enabled, any pending break exception is raised.

Use this function with scheme\_push\_break\_enable. If *post\_check* is non-zero and breaks are enabled after restoring the previous state, then any pending break exception is raised.

```
Scheme_Object*
scheme_current_continuation_marks(Scheme_Object* prompt_tag)
```

Like current-continuation-marks. Passing NULL as *prompt\_tag* is the same as providing the default continuation prompt tag.

## 8 Threads

The initializer function scheme\_basic\_env creates the main Scheme thread; all other threads are created through calls to scheme\_thread.

Information about each internal Scheme thread is kept in a Scheme\_Thread structure. A pointer to the current thread's structure is available as scheme\_current\_thread. A Scheme\_Thread structure includes the following fields:

- error\_buf the mz\_jmp\_buf value used to escape from errors. The error\_buf value of the current thread is available as scheme\_error\_buf.
- cjs.jumping\_to\_continuation a flag that distinguishes escaping-continuation invocations from error escapes. The cjs.jumping\_to\_continuation value of the current thread is available as scheme\_jumping\_to\_continuation.
- init\_config the thread's initial parameterization. See also §9 "Parameterizations".
- cell\_values The thread's values for thread cells (see also §9 "Parameterizations").
- next The next thread in the linked list of threads; this is NULL for the main thread.

The list of all scheduled threads is kept in a linked list; scheme\_first\_thread points to the first thread in the list. The last thread in the list is always the main thread.

### 8.1 Integration with Threads

Scheme's threads can break external C code under two circumstances:

- Pointers to stack-based values can be communicated between threads. For example, if thread A stores a pointer to a stack-based variable in a global variable, if thread B uses the pointer in the global variable, it may point to data that is not currently on the stack.
- C functions that can invoke Scheme (and also be invoked by Scheme) depend on strict function-call nesting. For example, suppose a function F uses an internal stack, pushing items on to the stack on entry and popping the same items on exit. Suppose also that F invokes Scheme to evaluate an expression. If the evaluation of this expression invokes F again in a new thread, but then returns to the first thread before completing the second F, then F's internal stack will be corrupted.

If either of these circumstances occurs, Scheme will probably crash.

# 8.2 Allowing Thread Switches

C code that performs substantial or unbounded work should occasionally call SCHEME\_USE\_FUEL—actually a macro—which allows Scheme to swap in another Scheme thread to run, and to check for breaks on the current thread. In particular, if breaks are enabled, then SCHEME\_USE\_FUEL may trigger an exception.

The macro consumes an integer argument. On most platforms, where thread scheduling is based on timer interrupts, the argument is ignored. On some platforms, however, the integer represents the amount of "fuel" that has been consumed since the last call to SCHEME\_USE\_FUEL. For example, the implementation of vector->list consumes a unit of fuel for each created cons cell:

```
Scheme_Object *scheme_vector_to_list(Scheme_Object *vec)
{
  int i;
  Scheme_Object *pair = scheme_null;

  i = SCHEME_VEC_SIZE(vec);

for (; i--; ) {
    SCHEME_USE_FUEL(1);
    pair = scheme_make_pair(SCHEME_VEC_ELS(vec)[i], pair);
  }

return pair;
}
```

The SCHEME\_USE\_FUEL macro expands to a C block, not an expression.

#### 8.3 Blocking the Current Thread

Embedding or extension code sometimes needs to block, but blocking should allow other Scheme threads to execute. To allow other threads to run, block using scheme\_block\_until. This procedure takes two functions: a polling function that tests whether the blocking operation can be completed, and a prepare-to-sleep function that sets bits in fd\_sets when Scheme decides to sleep (because all Scheme threads are blocked). Under Windows, an "fd\_set" can also accommodate OS-level semaphores or other handles via scheme\_add\_fd\_handle.

Since the functions passed to scheme\_block\_until are called by the Scheme thread scheduler, they must never raise exceptions, call scheme\_apply, or trigger the evaluation of Scheme code in any way. The scheme\_block\_until function itself may call the current exception handler, however, in reaction to a break (if breaks are enabled).

When a blocking operation is associated with an object, then the object might make sense as an argument to sync. To extend the set of objects accepted by sync, either register polling and sleeping functions with scheme\_add\_evt, or register a semaphore accessor with scheme\_add\_evt\_through\_sema.

The scheme\_signal\_received function can be called to wake up Scheme when it is sleeping. In particular, calling scheme\_signal\_received ensures that Scheme will poll all blocking synchronizations soon afterward. Furthermore, scheme\_signal\_received can be called from any OS-level thread. Thus, when no adequate prepare-to-sleep function can be implemented for scheme\_block\_until in terms of file descriptors or Windows handles, calling scheme\_signal\_received when the poll result changes will ensure that a poll is issued.

## 8.4 Threads in Embedded Scheme with Event Loops

When Scheme is embedded in an application with an event-based model (i.e., the execution of Scheme code in the main thread is repeatedly triggered by external events until the application exits) special hooks must be set to ensure that non-main threads execute correctly. For example, during the execution in the main thread, a new thread may be created; the new thread may still be running when the main thread returns to the event loop, and it may be arbitrarily long before the main thread continues from the event loop. Under such circumstances, the embedding program must explicitly allow Scheme to execute the non-main threads; this can be done by periodically calling the function scheme\_check\_threads.

Thread-checking only needs to be performed when non-main threads exist (or when there are active callback triggers). The embedding application can set the global function pointer scheme\_notify\_multithread to a function that takes an integer parameter and returns void. This function is be called with 1 when thread-checking becomes necessary, and then with 0 when thread checking is no longer necessary. An embedding program can use this information to prevent unnecessary scheme\_check\_threads polling.

The below code illustrates how MrEd formerly set up scheme\_check\_threads polling using the wxWindows wxTimer class. (Any regular event-loop-based callback is appropriate.) The scheme\_notify\_multithread pointer is set to MrEdInstallThreadTimer. (MrEd no longer work this way, however.)

```
class MrEdThreadTimer : public wxTimer
{
  public:
    void Notify(void); /* callback when timer expires */
};

static int threads_go;
static MrEdThreadTimer *theThreadTimer;
#define THREAD_WAIT_TIME 40
```

```
void MrEdThreadTimer::Notify()
{
   if (threads_go)
      Start(THREAD_WAIT_TIME, TRUE);

   scheme_check_threads();
}

static void MrEdInstallThreadTimer(int on)
{
   if (!theThreadTimer)
      theThreadTimer = new MrEdThreadTimer;

   if (on)
      theThreadTimer->Start(THREAD_WAIT_TIME, TRUE);
   else
      theThreadTimer->Stop();

   threads_go = on;
   if (on)
      do_this_time = 1;
}
```

An alternate architecture, which MrEd now uses, is to send the main thread into a loop, which blocks until an event is ready to handle. Scheme automatically takes care of running all threads, and it does so efficiently because the main thread blocks on a file descriptor, as explained in §8.3 "Blocking the Current Thread".

#### 8.4.1 Callbacks for Blocked Threads

Scheme threads are sometimes blocked on file descriptors, such as an input file or the X event socket. Blocked non-main threads do not block the main thread, and therefore do not affect the event loop, so scheme\_check\_threads is sufficient to implement this case correctly. However, it is wasteful to poll these descriptors with scheme\_check\_threads when nothing else is happening in the application and when a lower-level poll on the file descriptors can be installed. If the global function pointer scheme\_wakeup\_on\_input is set, then this case is handled more efficiently by turning off thread checking and issuing a "wakeup" request on the blocking file descriptors through scheme\_wakeup\_on\_input.

A scheme\_wakeup\_on\_input procedure takes a pointer to an array of three fd\_sets (sortof\footnote{To ensure maximum portability, use MZ\_FD\_XXX instead of FD\_XXX.}) and returns void. The scheme\_wakeup\_on\_input does not sleep; it just sets up callbacks on the specified file descriptors. When input is ready on any of those file descriptors, the callbacks are removed and scheme\_wake\_up is called.

For example, the X Windows version of MrEd formerly set  $scheme\_wakeup\_on\_input$  to this MrEdNeedWakeup:

```
static XtInputId *scheme_cb_ids = NULL;
static int num_cbs;
static void MrEdNeedWakeup(void *fds)
  int limit, count, i, p;
 fd_set *rd, *wr, *ex;
 rd = (fd_set *)fds;
 wr = ((fd_set *)fds) + 1;
 ex = ((fd_set *)fds) + 2;
 limit = getdtablesize();
 /* See if we need to do any work, really: */
 count = 0;
 for (i = 0; i < limit; i++) {
    if (MZ_FD_ISSET(i, rd))
     count++;
    if (MZ_FD_ISSET(i, wr))
      count++;
    if (MZ_FD_ISSET(i, ex))
     count++;
 if (!count)
   return;
 /* Remove old callbacks: */
 if (scheme_cb_ids)
    for (i = 0; i < num_cbs; i++)
     notify_set_input_func((Notify_client)NULL, (Notify_func)NULL,
                            scheme_cb_ids[i]);
 num_cbs = count;
 scheme_cb_ids = new int[num_cbs];
 /* Install callbacks */
 p = 0;
 for (i = 0; i < limit; i++) {
    if (MZ_FD_ISSET(i, rd))
     scheme_cb_ids[p++] = XtAppAddInput(wxAPP_CONTEXT, i,
                                          (XtPointer *)XtInputReadMask,
```

```
(XtInputCallbackProc)MrEdWakeUp, NULL);
    if (MZ_FD_ISSET(i, wr))
      scheme_cb_ids[p++] = XtAppAddInput(wxAPP_CONTEXT, i,
                                          (XtPointer *)XtInputWriteMask,
                                          (XtInputCallbackProc)MrEdWakeUp, NULL);
    if (MZ_FD_ISSET(i, ex))
      scheme_cb_ids[p++] = XtAppAddInput(wxAPP_CONTEXT, i,
                                          (XtPointer *)XtInputExceptMask,
                                          (XtInputCallbackProc)MrEdWakeUp,
                                          NULL);
/* callback function when input/exception is detected: */
Bool MrEdWakeUp(XtPointer, int *, XtInputId *)
  int i;
  if (scheme_cb_ids) {
    /* Remove all callbacks: */
    for (i = 0; i < num_cbs; i++)
     XtRemoveInput(scheme_cb_ids[i]);
    scheme_cb_ids = NULL;
    /* ''wake up'' */
    scheme_wake_up();
 return FALSE;
```

### 8.5 Sleeping by Embedded Scheme

When all Scheme threads are blocked, Scheme must "sleep" for a certain number of seconds or until external input appears on some file descriptor. Generally, sleeping should block the main event loop of the entire application. However, the way in which sleeping is performed may depend on the embedding application. The global function pointer scheme\_sleep can be set by an embedding application to implement a blocking sleep, although Scheme implements this function for you.

A scheme\_sleep function takes two arguments: a float and a void\*. The latter is really points to an array of three "fd\_set" records (one for read, one for write, and one for exceptions); these records are described further below. If the float argument is non-zero,

then the scheme\_sleep function blocks for the specified number of seconds, at most. The scheme\_sleep function should block until there is input one of the file descriptors specified in the "fd\_set," indefinitely if the float argument is zero.

The second argument to scheme\_sleep is conceptually an array of three fd\_set records, but always use scheme\_get\_fdset to get anything other than the zeroth element of this array, and manipulate each "fd\_set" with MZ\_FD\_XXX instead of FD\_XXX.

The following function mzsleep is an appropriate scheme\_sleep function for most any Unix or Windows application. (This is approximately the built-in sleep used by Scheme.)

```
void mzsleep(float v, void *fds)
{
   if (v) {
      sleep(v);
   } else {
      int limit;
      fd_set *rd, *wr, *ex;

# ifdef WIN32
      limit = 0;
# else
      limit = getdtablesize();
# endif

   rd = (fd_set *)fds;
   wr = (fd_set *)scheme_get_fdset(fds, 1);
   ex = (fd_set *)scheme_get_fdset(fds, 2);

   select(limit, rd, wr, ex, NULL);
   }
}
```

#### **8.6 Thread Functions**

Scheme\_Object\* scheme\_thread(Scheme\_Object\* thunk)

Creates a new thread, just like thread.

Like scheme\_thread, except that the created thread belongs to *cust* instead of the current custodian, it uses the given *config* for its initial configuration, it uses *cells* for its thread-cell table, and if *suspend\_to\_kill* is non-zero, then the thread is merely suspended when it would otherwise be killed (through either kill-thread or custodian-shutdown-all).

The *config* argument is typically obtained through scheme\_current\_config or scheme\_extend\_config. A *config* is immutable, so different threads can safely use the same value. The *cells* argument should be obtained from scheme\_inherit\_cells; it is mutable, and a particular cell table should be used by only one thread.

```
Scheme_Object* scheme_make_sema(long v)
```

Creates a new semaphore.

```
void scheme_post_sema(Scheme_Object* sema)
```

Posts to sema.

Waits on *sema*. If *try* is not 0, the wait can fail and 0 is returned for failure, otherwise 1 is returned.

```
void scheme_thread_block(float sleep_time)
```

Allows the current thread to be swapped out in favor of other threads. If *sleep\_time* positive, then the current thread will sleep for at least *sleep\_time* seconds.

After calling this function, a program should almost always call <code>scheme\_making\_progress</code> next. The exception is when <code>scheme\_thread\_block</code> is called in a polling loop that performs no work that affects the progress of other threads. In that case, <code>scheme\_making\_progress</code> should be called immediately after exiting the loop.

See also  $scheme_block_until$ , and see also the SCHEME\_USE\_FUEL macro in  $\S 8.2$  "Allowing Thread Switches".

Like scheme\_thread\_block, but breaks are enabled while blocking if break\_on is true.

```
void scheme_swap_thread(Scheme_Thread* thread)
```

Swaps out the current thread in favor of *thread*.

```
void scheme_break_thread(Scheme_Thread* thread)
```

Sends a break signal to the given thread.

```
int scheme_break_waiting(Scheme_Thread* thread)
```

Returns 1 if a break from break-thread or scheme\_break\_thread has occurred in the specified thread but has not yet been handled.

```
int scheme_block_until(Scheme_Ready_Fun f, Scheme_Needs_Wakeup_Fun fdf, Scheme_Object* data, float sleep)
```

The Scheme\_Ready\_Fun and Scheme\_Needs\_Wakeup\_Fun types are defined as follows:

Blocks the current thread until f with data returns a true value. The f function is called periodically—at least once per potential swap-in of the blocked thread—and it may be called multiple times even after it returns a true value. If f with data ever returns a true value, it must continue to return a true value until scheme\_block\_until returns. The argument to f is the same data as provided to scheme\_block\_until, and data is ignored otherwise. (The data argument is not actually required to be a Scheme\_Object\* value, because it is only used by f and fdf.)

If Scheme decides to sleep, then the *fdf* function is called to sets bits in *fds*, conceptually an array of three fd\_sets: one or reading, one for writing, and one for exceptions. Use scheme\_get\_fdset to get elements of this array, and manipulate an "fd\_set" with MZ\_FD\_XXX instead of FD\_XXX. Under Windows, an "fd\_set" can also accommodate OS-level semaphores or other handles via scheme\_add\_fd\_handle.

The *fdf* argument can be NULL, which implies that the thread becomes unblocked (i.e., *ready* changes its result to true) only through Scheme actions, and never through external processes (e.g., through a socket or OS-level semaphore)—with the exception that scheme\_signal\_received may be called to indicate an external change.

If sleep is a positive number, then  $scheme\_block\_until$  polls f at least every sleep seconds, but  $scheme\_block\_until$  does not return until f returns a true value. The call to  $scheme\_block\_until$  can return before sleep seconds if f returns a true value.

The return value from  $scheme\_block\_until$  is the return value of its most recent call to f, which enables f to return some information to the  $scheme\_block\_until$  caller.

See  $\S 8.3$  "Blocking the Current Thread" for information about restrictions on the f and fdf functions.

```
int scheme_block_until_enable_break(Scheme_Ready_Fun f, Scheme_Needs_Wakeup_Fun fdf, Scheme_Object* data, float sleep, int break\_on)
```

Like scheme\_block\_until, but breaks are enabled while blocking if *break\_on* is true.

Like scheme\_block\_until\_enable\_break, but the function returns if *unless\_evt* becomes ready, where *unless\_evt* is a port progress event implemented by scheme\_progress\_evt\_via\_get. See scheme\_make\_input\_port for more information.

```
void scheme_signal_received()
```

Indicates that an external event may have caused the result of a synchronization poll to have a different result. Unlike most other Scheme functions, this one can be called from any OS-level thread, and it wakes up if the Scheme thread if it is sleeping.

```
void scheme_check_threads()
```

This function is periodically called by the embedding program to give background processes time to execute. See §8.4 "Threads in Embedded Scheme with Event Loops" for more information.

```
void scheme_wake_up()
```

This function is called by the embedding program when there is input on an external file descriptor. See §8.5 "Sleeping by Embedded Scheme" for more information.

Extracts an "fd\_set" from an array passed to scheme\_sleep, a callback for scheme\_block\_until, or an input port callback for scheme\_make\_input\_port.

Adds an OS-level semaphore (Windows) or other waitable handle (Windows) to the "fd\_set" fds. When Scheme performs a "select" to sleep on fds, it also waits on the given semaphore or handle. This feature makes it possible for Scheme to sleep until it is awakened by an external process.

Scheme does not attempt to deallocate the given semaphore or handle, and the "select" call using fds may be unblocked due to some other file descriptor or handle in fds. If repost is a true value, then h must be an OS-level semaphore, and if the "select" unblocks due to a post on h, then h is reposted; this allows clients to treat fds-installed semaphores uniformly, whether or not a post on the semaphore was consumed by "select".

The scheme\_add\_fd\_handle function is useful for implementing the second procedure passed to scheme\_wait\_until, or for implementing a custom input port.

Under Unix and Mac OS X, this function has no effect.

Adds an OS-level event type (Windows) to the set of types in the "fd\_set" fds. When Scheme performs a "select" to sleep on fds, it also waits on events of them specified type. This feature makes it possible for Scheme to sleep until it is awakened by an external process.

The event mask is only used when some handle is installed with scheme\_add\_fd\_handle.

This awkward restriction may force you to create a dummy semaphore that is never posted.

Under Unix, and Mac OS X, this function has no effect.

```
void scheme_add_evt(Scheme_Type type,
Scheme_Ready_Fun ready,
Scheme_Needs_Wakeup_Fun wakeup,
Scheme_Wait_Filter_Fun filter,
int can_redirect)
```

The argument types are defined as follows:

Extends the set of waitable objects for sync to those with the type tag *type*. If *filter* is non-NULL, it constrains the new waitable set to those objects for which *filter* returns a non-zero value.

The *ready* and *wakeup* functions are used in the same way was the arguments to scheme block until.

The can\_redirect argument should be 0.

```
void scheme_add_evt_through_sema(Scheme_Type type,
Scheme_Wait_Sema_Fun getsema,
Scheme_Wait_Filter_Fun filter)
```

Like scheme\_add\_evt, but for objects where waiting is based on a semaphore. Instead of *ready* and *wakeup* functions, the *getsema* function extracts a semaphore for a given object:

If a successful wait should leave the semaphore waited, then *getsema* should set \**repost* to 0. Otherwise, the given semaphore will be re-posted after a successful wait. A *getsema* function should almost always set \**repost* to 1.

```
void scheme_making_progress()
```

Notifies the scheduler that the current thread is not simply calling scheme\_thread\_block in a loop, but that it is actually making progress.

```
int scheme_tls_allocate()
```

Allocates a thread local storage index to be used with scheme\_tls\_set and scheme\_tls\_get.

Stores a thread-specific value using an index allocated with scheme\_tls\_allocate.

```
void* scheme_tls_get(int index)
```

Retrieves a thread-specific value installed with scheme\_tls\_set. If no thread-specific value is available for the given index, NULL is returned.

Calls *prim* with the given *argc* and *argv* with breaks enabled. The *prim* function can block, in which case it might be interrupted by a break. The *prim* function should not block, yield, or check for breaks after it succeeds, where "succeeds" depends on the operation. For example, tcp-accept/enable-break is implemented by wrapping this function around the implementation of tcp-accept; the tcp-accept implementation does not block or yield after it accepts a connection.

Prevents Scheme thread swaps until scheme\_end\_atomic or scheme\_end\_atomic\_no\_swap is called. Start-atomic and end-atomic pairs can be nested.

```
void scheme_end_atomic()
```

Ends an atomic region with respect to Scheme threads. The current thread may be swapped

out immediately (i.e., the call to scheme\_end\_atomic is assumed to be a safe point for thread swaps).

```
void scheme_end_atomic_no_swap()
```

Ends an atomic region with respect to Scheme threads, and also prevents an immediate thread swap. (In other words, no Scheme thread swaps will occur until a future safe point.)

```
void scheme_add_swap_callback(Scheme_Closure_Func f, Scheme_Object* data)
```

Registers a callback to be invoked just after a Scheme thread is swapped in. The *data* is provided back to *f* when it is called, where Closure\_Func is defined as follows:

```
typedef Scheme_Object *(*Scheme_Closure_Func)(Scheme_Object *);
```

Like scheme\_add\_swap\_callback, but registers a callback to be invoked just before a Scheme thread is swapped out.

#### 9 Parameterizations

A *parameterization* is a set of parameter values. Each thread has its own initial parameterization, which is extended functionally and superseded by parameterizations that are attached to a particular continuation mark.

Parameterization information is stored in a Scheme\_Config record. For the currently executing thread, scheme\_current\_config returns the current parameterization.

To obtain parameter values, a Scheme\_Config is combined with the current threads Scheme\_Thread\_Cell\_Table, as stored in the thread record's cell\_values field.

Parameter values for built-in parameters are obtained and modified (for the current thread) using scheme\_get\_param and scheme\_set\_param. Each parameter is stored as a Scheme\_Object \* value, and the built-in parameters are accessed through the following indices:

- MZCONFIG\_ENV current-namespace (use scheme\_get\_env)
- MZCONFIG\_INPUT\_PORT current-input-port
- MZCONFIG\_OUTPUT\_PORT current-output-port
- MZCONFIG\_ERROR\_PORT current-error-port
- MZCONFIG\_ERROR\_DISPLAY\_HANDLER error-display-handler
- MZCONFIG\_ERROR\_PRINT\_VALUE\_HANDLER error-value->string-handler
- MZCONFIG\_EXIT\_HANDLER exit-handler
- MZCONFIG\_INIT\_EXN\_HANDLER uncaught-exception-handler
- MZCONFIG\_EVAL\_HANDLER current-eval
- MZCONFIG\_LOAD\_HANDLER current-load
- MZCONFIG\_PRINT\_HANDLER current-print
- MZCONFIG\_PROMPT\_READ\_HANDLER current-prompt-read
- MZCONFIG\_CAN\_READ\_GRAPH read-accept-graph
- MZCONFIG\_CAN\_READ\_COMPILED read-accept-compiled
- MZCONFIG\_CAN\_READ\_BOX read-accept-box
- MZCONFIG\_CAN\_READ\_PIPE\_QUOTE read-accept-bar-quote
- MZCONFIG\_PRINT\_GRAPH print-graph

- MZCONFIG\_PRINT\_STRUCT print-struct
- MZCONFIG\_PRINT\_BOX print-box
- MZCONFIG\_CASE\_SENS read-case-sensitive
- MZCONFIG\_SQUARE\_BRACKETS\_ARE\_PARENS read-square-brackets-asparens
- MZCONFIG\_CURLY\_BRACES\_ARE\_PARENS read-curly-braces-as-parens
- MZCONFIG\_ERROR\_PRINT\_WIDTH error-print-width
- MZCONFIG\_ALLOW\_SET\_UNDEFINED allow-compile-set!-undefined
- MZCONFIG\_CUSTODIAN current-custodian
- MZCONFIG\_USE\_COMPILED\_KIND use-compiled-file-paths
- MZCONFIG\_LOAD\_DIRECTORY current-load-relative-directory
- MZCONFIG\_COLLECTION\_PATHS current-library-collection-paths
- MZCONFIG\_PORT\_PRINT\_HANDLER global-port-print-handler
- MZCONFIG\_LOAD\_EXTENSION\_HANDLER current-load-extension

To get or set a parameter value for a thread other than the current one, use scheme\_get\_thread\_param and scheme\_set\_thread\_param, each of which takes a Scheme\_Thread\_Cell\_Table to use in resolving or setting a parameter value.

When installing a new parameter with scheme\_set\_param, no check is performed on the supplied value to ensure that it is a legal value for the parameter; this is the responsibility of the caller of scheme\_set\_param. Note that Boolean parameters should only be set to the values #t and #f.

New primitive parameter indices are created with scheme\_new\_param and implemented with scheme\_make\_parameter and scheme\_param\_config.

Gets the current value (for the current thread) of the parameter specified by param\_id.

Sets the current value (for the current thread) of the parameter specified by *param\_id*.

Like scheme\_get\_param, but using an arbitrary thread's cell-value table.

Like scheme\_set\_param, but using an arbitrary thread's cell-value table.

Creates and returns a parameterization that extends *base* with a new value *v* (in all threads) for the parameter *param\_id*. Use scheme\_install\_config to make this configuration active in the current thread.

```
void scheme_install_config(Scheme_Config* config)
```

Adjusts the current thread's continuation marks to make *config* the current parameterization. Typically, this function is called after scheme\_push\_continuation\_frame to establish a new continuation frame, and then scheme\_pop\_continuation\_frame is called later to remove the frame (and thus the parameterization).

```
Scheme_Thread_Cell_Table*
scheme_inherit_cells(Scheme_Thread_Cell_Table* cells)
```

Creates a new thread-cell-value table, copying values for preserved thread cells from cells.

```
int scheme_new_param()
```

Allocates a new primitive parameter index. This function must be called *before* scheme\_basic\_env, so it is only available to embedding applications (i.e., not extensions).

Use this function instead of the other primitive-constructing functions, like scheme\_make\_prim, to create a primitive parameter procedure. See also scheme\_param\_config, below. This function is only available to embedding applications (i.e., not extensions).

Call this procedure in a primitive parameter procedure to implement the work of getting or setting the parameter. The *name* argument should be the parameter procedure name; it is used to report errors. The *param* argument is a fixnum corresponding to the primitive parameter index returned by scheme\_new\_param. The *argc* and *argv* arguments should be the un-touched and un-tested arguments that were passed to the primitive parameter. Argument-checking is performed within scheme\_param\_config using *arity*, *check*, *expected*, and *is-bool*:

- If *arity* is non-negative, potential parameter values must be able to accept the specified number of arguments. The *check* and *expected* arguments should be NULL.
- If *check* is not NULL, it is called to check a potential parameter value. The arguments passed to *check* are always 1 and an array that contains the potential parameter value. If *isbool* is 0 and *check* returns scheme\_false, then a type error is reported using *name* and *expected*. If *isbool* is 1, then a type error is reported only when *check* returns NULL and any non-NULL return value is used as the actual value to be stored for the parameter.
- Otherwise, isbool should be 1. A potential procedure argument is then treated as a Boolean value.

This function is only available to embedding applications (i.e., not extensions).

# 10 Continuation Marks

A mark can be attached to the current continuation frame using scheme\_set\_cont\_mark. To force the creation of a new frame (e.g., during a nested function call within your function), use scheme\_push\_continuation\_frame, and then remove the frame with scheme\_pop\_continuation\_frame.

Add/sets a continuation mark in the current continuation.

void scheme\_push\_continuation\_frame(Scheme\_Cont\_Frame\_Data\* data)

Creates a new continuation frame. The *data* record need not be initialized, and it can be allocated on the C stack. Supply *data* to scheme\_pop\_continuation\_frame to remove the continuation frame.

void scheme\_pop\_continuation\_frame(Scheme\_Cont\_Frame\_Data\* data)

Removes a continuation frame created by scheme\_pop\_continuation\_frame.

## 11 String Encodings

The scheme\_utf8\_decode function decodes a char array as UTF-8 into either a UCS-4 mzchar array or a UTF-16 short array. The scheme\_utf8\_encode function encodes either a UCS-4 mzchar array or a UTF-16 short array into a UTF-8 char array.

These functions can be used to check or measure an encoding or decoding without actually producing the result decoding or encoding, and variations of the function provide control over the handling of decoding errors.

Decodes a byte array as UTF-8 to produce either Unicode code points into *us* (when *utf16* is zero) or UTF-16 code units into *us* cast to short\* (when *utf16* is non-zero). No nul terminator is added to *us*.

The result is non-negative when all of the given bytes are decoded, and the result is the length of the decoding (in mzchars or shorts). A -2 result indicates an invalid encoding sequence in the given bytes (possibly because the range to decode ended mid-encoding), and a -3 result indicates that decoding stopped because not enough room was available in the result string.

The *start* and *end* arguments specify a range of s to be decoded. If *end* is negative, strlen(s) is used as the end.

If us is NULL, then decoded bytes are not produced, but the result is valid as if decoded bytes were written. The dstart and dend arguments specify a target range in us (in mzchar or short units) for the decoding; a negative value for dend indicates that any number of bytes can be written to us, which is normally sensible only when us is NULL for measuring the length of the decoding.

If ipos is non-NULL, it is filled with the first undecoded index within s. If the function result is non-negative, then \*ipos is set to the ending index (with is end if non-negative, strlen(s) otherwise). If the result is -1 or -2, then \*ipos effectively indicates how many bytes were decoded before decoding stopped.

If permissive is non-zero, it is used as the decoding of bytes that are not part of a valid UTF-8

encoding. Thus, the function result can be -2 only if *permissive* is 0.

This function does not allocate or trigger garbage collection.

Like scheme\_utf8\_decode, but the result is always the number of the decoded mzchars or shorts. If a decoding error is encountered, the result is still the size of the decoding up until the error.

Like scheme\_utf8\_decode, but with fewer arguments. The decoding produces UCS-4 mzchars. If the buffer us is non-NULL, it is assumed to be long enough to hold the decoding (which cannot be longer than the length of the input, though it may be shorter). If len is negative, strlen(s) is used as the input length.

Like scheme\_utf8\_decode, but with fewer arguments. The decoding produces UCS-4 mzchars. If the buffer *us* must be non-NULL, and it is assumed to be long enough to hold the decoding (which cannot be longer than the length of the input, though it may be shorter). If *len* is negative, strlen(s) is used as the input length.

In addition to the result of scheme\_utf8\_decode, the result can be -1 to indicate that the input ended with a partial (valid) encoding. A -1 result is possible even when *permissive* is non-zero.

```
\begin{tabular}{ll} \verb|mzchar*| & scheme_utf8_decode_to_buffer(const unsigned char*| s, \\ & int len, \\ & mzchar*| buf, \\ & int blen) \end{tabular}
```

Like scheme\_utf8\_decode\_all with *permissive* as 0, but if *buf* is not large enough (as indicated by *blen*) to hold the result, a new buffer is allocated. Unlike other functions, this one adds a nul terminator to the decoding result. The function result is either *buf* (if it was big enough) or a buffer allocated with scheme\_malloc\_atomic.

Like scheme\_utf8\_decode\_to\_buffer, but the length of the result (not including the terminator) is placed into *ulen* if *ulen* is non-NULL.

Like scheme\_utf8\_decode, but without producing the decoded mzchars, and always returning the number of decoded mzchars up until a decoding error (if any). If *might\_continue* is non-zero, the a partial valid encoding at the end of the input is not decoded when *permissive* is also non-zero.

If *state* is non-NULL, it holds information about partial encodings; it should be set to zero for an initial call, and then passed back to scheme\_utf8\_decode along with bytes that extend the given input (i.e., without any unused partial encodings). Typically, this mode makes sense only when *might\_continue* and *permissive* are non-zero.

Encodes the given UCS-4 array of mzchars (if utf16 is zero) or UTF-16 array of shorts (if

utf16 is non-zero) into s. The end argument must be no less than start.

The array *s* is assumed to be long enough to contain the encoding, but no encoding is written if *s* is NULL. The *dstart* argument indicates a starting place in *s* to hold the encoding. No nul terminator is added to *s*.

The result is the number of bytes produced for the encoding (or that would be produced if *s* was non-NULL). Encoding never fails.

This function does not allocate or trigger garbage collection.

Like scheme\_utf8\_encode with 0 for start, len for end, 0 for dstart and 0 for utf16.

Like scheme\_utf8\_encode\_all, but the length of *buf* is given, and if it is not long enough to hold the encoding, a buffer is allocated. A nul terminator is added to the encoded array. The result is either *buf* or an array allocated with scheme\_malloc\_atomic.

Like scheme\_utf8\_encode\_to\_buffer, but the length of the resulting encoding (not including a nul terminator) is reported in *rlen* if it is non-NULL.

Converts a UCS-4 encoding (the indicated range of *text*) to a UTF-16 encoding. The *end* argument must be no less than *start*.

A result buffer is allocated if *buf* is not long enough (as indicated by *bufsize*). If *ulen* is non-NULL, it is filled with the length of the UTF-16 encoding. The *term\_size* argument indicates a number of shorts to reserve at the end of the result buffer for a terminator (but no terminator is actually written).

Converts a UTF-16 encoding (the indicated range of *text*) to a UCS-4 encoding. The *end* argument must be no less than *start*.

A result buffer is allocated if *buf* is not long enough (as indicated by *bufsize*). If *ulen* is non-NULL, it is filled with the length of the UCS-4 encoding. The *term\_size* argument indicates a number of mzchars to reserve at the end of the result buffer for a terminator (but no terminator is actually written).

## 12 Bignums, Rationals, and Complex Numbers

Scheme supports integers of an arbitrary magnitude; when an integer cannot be represented as a fixnum (i.e., 30 or 62 bits plus a sign bit), then it is represented by the Scheme type scheme\_bignum\_type. There is no overlap in integer values represented by fixnums and bignums.

Rationals are implemented by the type scheme\_rational\_type, composed of a numerator and a denominator. The numerator and denominator fixnums or bignums (possibly mixed).

Complex numbers are implemented by the types scheme\_complex\_type and scheme\_complex\_izi\_type, composed of a real and imaginary part. The real and imaginary parts will either be both flonums, both exact numbers (fixnums, bignums, and rationals can be mixed in any way), or one part will be exact 0 and the other part will be a flonum. If the inexact part is inexact 0, the type is scheme\_complex\_izi\_type, otherwise the type is scheme\_complex\_type; this distinction make it easy to test whether a complex number should be treated as a real number.

```
int scheme_is_exact(Scheme_Object* n)
```

Returns 1 if n is an exact number, 0 otherwise (n need not be a number).

```
int scheme_is_inexact(Scheme_Object* n)
```

Returns 1 if n is an inexact number, 0 otherwise (n need not be a number).

```
Scheme_Object* scheme_make_bignum(long v)
```

Creates a bignum representing the integer  $\nu$ . This can create a bignum that otherwise fits into a fixnum. This must only be used to create temporary values for use with the bignum functions. Final results can be normalized with scheme\_bignum\_normalize. Only normalized numbers can be used with procedures that are not specific to bignums.

```
Scheme_Object* scheme_make_bignum_from_unsigned(unsigned long v)
```

Like scheme\_make\_bignum, but works on unsigned integers.

```
double scheme_bignum_to_double(Scheme_Object* n)
```

Converts a bignum to a floating-point number, with reasonable but unspecified accuracy.

```
float scheme_bignum_to_float(Scheme_Object* n)
```

If PLT Scheme is not compiled with single-precision floats, this procedure is actually a macro alias for scheme\_bignum\_to\_double.

```
Scheme_Object* scheme_bignum_from_double(double d)
```

Creates a bignum that is close in magnitude to the floating-point number d. The conversion accuracy is reasonable but unspecified.

```
Scheme_Object* scheme_bignum_from_float(float f)
```

If PLT Scheme is not compiled with single-precision floats, this procedure is actually a macro alias for scheme\_bignum\_from\_double.

Writes a bignum into a newly allocated byte string.

Reads a bignum from a mzchar string, starting from position *offset* in *str*. If the string does not represent an integer, then NULL will be returned. If the string represents a number that fits in 31 bits, then a scheme\_integer\_type object will be returned.

```
Scheme_Object* scheme_read_bignum_bytes(char* str, int offset, int radix)
```

Like scheme\_read\_bignum, but from a UTF-8-encoding byte string.

```
Scheme_Object* scheme_bignum_normalize(Scheme_Object* n)
```

If n fits in 31 bits, then a scheme\_integer\_type object will be returned. Otherwise, n is returned.

```
Scheme_Object* scheme_make_rational(Scheme_Object* n, Scheme_Object* d)
```

Creates a rational from a numerator and denominator. The n and d parameters must be fixnums or bignums (possibly mixed). The resulting will be normalized (thus, a bignum or fixnum might be returned).

double scheme\_rational\_to\_double(Scheme\_Object\* n)

Converts the rational n to a double.

float scheme\_rational\_to\_float(Scheme\_Object\* n)

If PLT Scheme is not compiled with single-precision floats, this procedure is actually a macro alias for scheme\_rational\_to\_double.

Scheme\_Object\* scheme\_rational\_numerator(Scheme\_Object\* n)

Returns the numerator of the rational n.

Scheme\_Object\* scheme\_rational\_denominator(Scheme\_Object\* n)

Returns the denominator of the rational n.

Scheme\_Object\* scheme\_rational\_from\_double(double d)

Converts the given double into a maximally-precise rational.

Scheme\_Object\* scheme\_rational\_from\_float(float d)

If PLT Scheme is not compiled with single-precision floats, this procedure is actually a macro alias for scheme\_rational\_from\_double.

Creates a complex number from real and imaginary parts. The r and i arguments must be fixnums, bignums, flonums, or rationals (possibly mixed). The resulting number will be normalized (thus, a real number might be returned).

Scheme\_Object\* scheme\_complex\_real\_part(Scheme\_Object\* n)

Returns the real part of the complex number n.

Scheme\_Object\* scheme\_complex\_imaginary\_part(Scheme\_Object\* n)

Returns the imaginary part of the complex number n.

## 13 Ports and the Filesystem

Ports are represented as Scheme values with the types scheme\_input\_port\_type and scheme\_output\_port\_type. The function scheme\_read takes an input port value and returns the next S-expression from the port. The function scheme\_write takes an output port and a value and writes the value to the port. Other standard low-level port functions are also provided, such as scheme\_getc.

File ports are created with scheme\_make\_file\_input\_port and scheme\_make\_file\_output\_port; these functions take a FILE \* file pointer and return a Scheme port. Strings are read or written with scheme\_make\_byte\_string\_input\_port, which takes a nul-terminated byte string, and scheme\_make\_byte\_string\_output\_port, which takes no arguments. The contents of a string output port are obtained with scheme\_get\_byte\_string\_output.

Custom ports, with arbitrary read/write handlers, are created with scheme\_make\_input\_port and scheme\_make\_output\_port.

When opening a file for any reason using a name provided from Scheme, use scheme\_expand\_filename to normalize the filename and resolve relative paths.

```
Scheme_Object* scheme_read(Scheme_Object* port)
```

reads the next S-expression from the given input port.

writes the Scheme value *obj* to the given output port.

Like scheme\_write, but the printing is truncated to *n* bytes. (If printing is truncated, the last bytes are printed as ".".)

displays the Scheme value *obj* to the given output port.

Like scheme\_display, but the printing is truncated to *n* bytes. (If printing is truncated, the last three bytes are printed as ".".)

Writes *len* bytes of *str* to the given output port.

Writes *len* characters of *str* to the given output port.

Writes *len* bytes of *str*, starting with the *d*th character. Bytes are written to the given output port, and errors are reported as from *who*.

If *rarely\_block* is 0, the write blocks until all *len* bytes are written, possibly to an internal buffer. If *rarely\_block* is 2, the write never blocks, and written bytes are not buffered. If *rarely\_block* is 1, the write blocks only until at least one byte is written (without buffering) or until part of an internal buffer is flushed.

Supplying 0 for *len* corresponds to a buffer-flush request. If *rarely\_block* is 2, the flush request is non-blocking, and if *rarely\_block* is 0, it is blocking. (A *rarely\_block* of 1 is the same as 0 in this case.)

The result is -1 if no bytes are written from *str* and unflushed bytes remain in the internal buffer. Otherwise, the return value is the number of written characters.

Like scheme\_put\_byte\_string, but for a mzchar string, and without the non-blocking option.

Prints the Scheme value *obj* using write to a newly allocated string. If *len* is not NULL, \**len* is set to the length of the bytes string.

Like scheme\_write\_to\_string, but the string is truncated to n bytes. (If the string is truncated, the last three bytes are ".".)

```
char* scheme_display_to_string(Scheme_Object* obj, long* len)
```

Prints the Scheme value *obj* using display to a newly allocated string. If *len* is not NULL, \**len* is set to the length of the string.

Like  $scheme_display_to_string$ , but the string is truncated to n bytes. (If the string is truncated, the last three bytes are ".".)

```
void scheme_debug_print(Scheme_Object* obj)
```

Prints the Scheme value *obj* using write to the main thread's output port.

```
void scheme_flush_output(Scheme_Object* port)
```

If *port* is a file port, a buffered data is written to the file. Otherwise, there is no effect. *port* must be an output port.

```
int scheme_get_byte(Scheme_Object* port)
```

Get the next byte from the given input port. The result can be EOF.

```
int scheme_getc(Scheme_Object* port)
```

Get the next character from the given input port (by decoding bytes as UTF-8). The result can be E0F.

```
int scheme_peek_byte(Scheme_Object* port)
```

Peeks the next byte from the given input port. The result can be EOF.

```
int scheme_peekc(Scheme_Object* port)
```

Peeks the next character from the given input port (by decoding bytes as UTF-8). The result can be E0F.

Like scheme\_peek\_byte, but with a skip count. The result can be EOF.

Like scheme\_peekc, but with a skip count. The result can be EOF.

Gets multiple bytes at once from a port, reporting errors with the name *who*. The *size* argument indicates the number of requested bytes, to be put into the *buffer* array starting at *offset*. The return value is the number of bytes actually read, or EOF if an end-of-file is encountered without reading any bytes.

If *only\_avail* is 0, then the function blocks until *size* bytes are read or an end-of-file is reached. If *only\_avail* is 1, the function blocks only until at least one byte is read. If *only\_avail* is 2, the function never blocks. If *only\_avail* is -1, the function blocks only until at least one byte is read but also allows breaks (with the guarantee that bytes are read or a break is raised, but not both).

If *peek* is non-zero, then the port is peeked instead of read. The *peek\_skip* argument indicates a portion of the input stream to skip as a non-negative, exact integer (fixnum or bignum). In this case, an *only\_avail* value of 1 means to continue the skip until at least one byte can be returned, even if it means multiple blocking reads to skip bytes.

If peek is zero, then peek\_skip should be either NULL (which means zero) or the fixnum zero.

Like scheme\_get\_byte\_string, but for characters (by decoding bytes as UTF-8), and without the non-blocking option.

For backward compatibility: calls scheme\_get\_byte\_string in essentially the obvious way with *only\_avail* as 0; if *size* is negative, then it reads *-size* bytes with *only\_avail* as 1.

Puts the byte *ch* back as the next character to be read from the given input port. The character need not have been read from *port*, and scheme\_ungetc can be called to insert up to five characters at the start of *port*.

Use scheme\_get\_byte followed by scheme\_ungetc only when your program will certainly call scheme\_get\_byte again to consume the byte. Otherwise, use scheme\_peek\_byte, because some a port may implement peeking and getting differently.

```
int scheme_byte_ready(Scheme_Object* port)
```

Returns 1 if a call to scheme\_get\_byte is guaranteed not to block for the given input port.

```
int scheme_char_ready(Scheme_Object* port)
```

Returns 1 if a call to scheme\_getc is guaranteed not to block for the given input port.

Requests that appropriate bits are set in *fds* to specify which file descriptors(s) the given input port reads from. (*fds* is sort a pointer to an fd\_set struct; see §8.4.1 "Callbacks for Blocked Threads".)

```
long scheme_tell(Scheme_Object* port)
```

Returns the current read position of the given input port, or the current file position of the given output port.

```
long scheme_tell_line(Scheme_Object* port)
```

Returns the current read line of the given input port. If lines are not counted, -1 is returned.

```
void scheme_count_lines(Scheme_Object* port)
```

Turns on line-counting for the given input port. To get accurate line counts, call this function immediately after creating a port.

Sets the file position of the given input or output port (from the start of the file). If the port does not support position setting, an exception is raised.

```
void scheme_close_input_port(Scheme_Object* port)
```

Closes the given input port.

```
void scheme_close_output_port(Scheme_Object* port)
```

Closes the given output port.

Fills \*fd with a file-descriptor value for port if one is available (i.e., the port is a file-stream port and it is not closed). The result is non-zero if the file-descriptor value is available, zero otherwise. Under Windows, a "file dscriptor" is a file HANDLE.

Fills \*s with a socket value for *port* if one is available (i.e., the port is a TCP port and it is not closed). The result is non-zero if the socket value is available, zero otherwise. Under Windows, a socket value has type SOCKET.

```
Scheme_Object* scheme_make_port_type(char* name)
```

Creates a new port subtype.

Creates a new input port with arbitrary control functions. The *subtype* is an arbitrary value to distinguish the port's class. The pointer *data* will be installed as the port's user data, which can be extracted/set with the SCHEME\_INPORT\_VAL macro. The *name* object is used as the port's name (for object-name and as the default source name for read-syntax).

If *must\_close* is non-zero, the new port will be registered with the current custodian, and *close\_fun* is guaranteed to be called before the port is garbage-collected.

Although the return type of scheme\_make\_input\_port is Scheme\_Input\_Port\*, it can be cast into a Scheme\_Object\*.

The functions are as follows.

Reads bytes into *buffer*, starting from *offset*, up to *size* bytes (i.e., *buffer* is at least *offset* plus *size* long). If *nonblock* is 0, then the function can block indefinitely, but it should return when at least one byte of data is available. If *nonblock* is 1, the function should never block. If *nonblock* is 2, a port in unbuffered mode should return only bytes previously forced to be buffered; other ports should treat a *nonblock* of 2 like 1. If *nonblock* is -1, the function can block, but should enable breaks while blocking. The function should return 0 if no bytes are ready in non-blocking mode. It should return EOF if an end-of-file is reached (and no bytes were read into *buffer*). Otherwise, the function should return the number of read bytes. The function can raise an exception to report an error.

The *unless* argument will be non-NULL only when *nonblocking* is non-zero (except as noted below), and only if the port supports progress events. If *unless* is non-NULL and SCHEME\_CDR(*unless*) is non-NULL, the latter is a progress event specific to the port. The *get\_bytes\_fun* function should return SCHEME\_UNLESS\_READY instead of reading bytes if the event in *unless* becomes ready before bytes can be read. In particular, *get\_bytes\_fun* should check the event in *unless* after any operation that may allow Scheme thread swaps. If the read must block, then it should unblock if the event in *unless* becomes ready.

If scheme\_progress\_evt\_via\_get is used for progress\_evt\_fun, then unless can be non-NULL even when nonblocking is 0. In all modes, get\_bytes\_fun must call scheme\_unless\_ready to check unless\_evt. Furthermore, after any potentially thread-swapping operation, get\_bytes\_fun must call scheme\_wait\_input\_allowed, because another thread may be attempting to commit, and unless\_evt must be checked after scheme\_wait\_input\_allowed returns. To block, the port should use scheme\_block\_until\_unless instead of scheme\_block\_until. Finally, in blocking mode, get\_bytes\_fun must return after immediately reading data, without allowing a Scheme thread swap.

Can be NULL to use a default implementation of peeking that uses *get\_bytes\_fun*. Otherwise, the protocol is the same as for *get\_bytes\_fun*, except that an extra *skip* argument indicates the number of input elements to skip (but *skip* does not apply to *buffer*). The *skip* value will be a non-negative exact integer, either a fixnum or a bignum.

```
Scheme_Object* progress_evt_fun(Scheme_Input_Port* port)
```

Called to obtain a progress event for the port, such as for port-progress-evt. This function can be NULL if the port does not support progress events. Use progress\_evt\_via\_get to obtain a default implementation, in which case <code>peeked\_read\_fun</code> should be <code>peeked\_read\_via\_get</code>, and <code>get\_bytes\_fun</code> and <code>peek\_bytes\_fun</code> should handle <code>unless</code> as described above.

Called to commit previously peeked bytes, just like the sixth argument to make-input-port. Use peeked\_read\_via\_get for the default implementation of commits when *progress\_evt\_fun* is progress\_evt\_via\_get.

```
int char_ready_fun(Scheme_Input_Port* port)
```

Returns 1 when a non-blocking *get\_bytes\_fun* will return bytes or an EOF.

```
void close_fun(Scheme_Input_Port* port)
```

Called to close the port. The port is not considered closed until the function returns.

Called when the port is blocked on a read; <code>need\_wakeup\_fun</code> should set appropriate bits in <code>fds</code> to specify which file descriptor(s) it is blocked on. The <code>fds</code> argument is conceptually an array of three <code>fd\_set</code> structs (one for read, one for write, one for exceptions), but manipulate this array using <code>scheme\_get\_fdset</code> to get a particular element of the array, and use <code>MZ\_FD\_XXX</code> instead of <code>FD\_XXX</code> to manipulate a single "fd\_set". Under Windows, the first "fd\_set" can also contain OS-level semaphores or other handles via <code>scheme\_add\_fd\_handle</code>.

Creates a new output port with arbitrary control functions. The *subtype* is an arbitrary value to distinguish the port's class. The pointer *data* will be installed as the port's user data, which can be extracted/set with the SCHEME\_OUTPORT\_VAL macro. The *name* object is used as the port's name.

If *must\_close* is non-zero, the new port will be registered with the current custodian, and *close\_fun* is guaranteed to be called before the port is garbage-collected.

Although the return type of scheme\_make\_output\_port is Scheme\_Output\_Port\*, it can be cast into a Scheme\_Object\*.

The functions are as follows.

Returns an event that writes up to *size* bytes atomically when event is chosen in a synchronization. Supply NULL if bytes cannot be written atomically, or supply scheme\_write\_evt\_via\_write to use the default implementation in terms of write\_bytes\_fun (with *rarely\_block* as 2).

Write bytes from *buffer*, starting from *offset*, up to *size* bytes (i.e., *buffer* is at least *offset* plus *size* long). If *rarely\_block* is 0, then the function can block indefinitely, and it can buffer output. If *rarely\_block* is 2, the function should never block, and it should not buffer output. If *rarely\_block* is 1, the function should not buffer data, and it should block only until writing at least one byte, either from *buffer* or an internal buffer. The function should return the number of bytes from *buffer* that were written; when *rarely\_block* is non-zero and bytes remain in an internal buffer, it should return -1. The *size* argument can be 0 when *rarely\_block* is 0 for a blocking flush, and it can be 0 if *rarely\_block* is 2 for a non-blocking flush. If *enable\_break* is true, then it should enable breaks while blocking. The function can raise an exception to report an error.

```
int char_ready_fun(Scheme_Output_Port* port)
```

Returns 1 when a non-blocking *write\_bytes\_fun* will write at least one byte or flush at least one byte from the port's internal buffer.

```
void close_fun(Scheme_Output_Port* port)
```

Called to close the port. The port is not considered closed until the function returns. This function is allowed to block (usually to flush a buffer) unless scheme\_close\_should\_force\_port\_closed returns a non-zero result, in which case the function must return without blocking.

Called when the port is blocked on a write; *need\_wakeup\_fun* should set appropriate bits in *fds* to specify which file descriptor(s) it is blocked on. The *fds* 

argument is conceptually an array of three fd\_set structs (one for read, one for write, one for exceptions), but manipulate this array using scheme\_get\_fdset to get a particular element of the array, and use MZ\_FD\_XXX instead of FD\_XXX to manipulate a single "fd\_set". Under Windows, the first "fd\_set" can also contain OS-level semaphores or other handles via scheme\_add\_fd\_handle.

Returns an event that writes v atomically when event is chosen in a synchronization. Supply NULL if specials cannot be written atomically (or at all), or supply scheme\_write\_special\_evt\_via\_write\_special to use the default implementation in terms of write\_special\_fun (with  $non\_block$  as 1).

Called to write the special value v for write-special (when  $non\_block$  is 0) or write-special-avail\* (when  $non\_block$  is 1). If NULL is supplied instead of a function pointer, then write-special and write-special-avail\* produce an error for this port.

```
Scheme_Object* scheme_make_file_input_port(FILE* fp)
```

Creates a Scheme input file port from an ANSI C file pointer. The file must never block on reads.

```
Scheme_Object* scheme_open_input_file(const char* filename, const char* who)
```

Opens *filename* for reading. In an exception is raised, the exception message uses *who* as the name of procedure that raised the exception.

Creates a Scheme input file port from an ANSI C file pointer. The file must never block on reads. The *name* argument is used as the port's name.

Opens *filename* for writing in 'truncate/replace mode. If an exception is raised, the exception message uses *who* as the name of procedure that raised the exception.

```
Scheme_Object* scheme_make_file_output_port(FILE* fp)
```

Creates a Scheme output file port from an ANSI C file pointer. The file must never block on writes.

```
Scheme_Object* scheme_make_fd_input_port(int fd,

Scheme_Object* name,
int regfile,
int win_textmode,
int fd,
Scheme_Object* name,
int regfile,
int win_textmode,
int read_too,
long s,
const char* name,
int close,
Scheme_Object** inp,
Scheme_Object** outp,
char* str)
```

Creates a Scheme input port from a byte string; successive read-chars on the port return successive bytes in the string.

```
Scheme_Object* scheme_make_byte_string_output_port()
```

Creates a Scheme output port; all writes to the port are kept in a byte string, which can be obtained with scheme\_get\_byte\_string\_output.

```
char* scheme_get_byte_string_output(Scheme_Object* port)
```

Returns (in a newly allocated byte string) all data that has been written to the given string output port so far. (The returned string is nul-terminated.)

Returns (in a newly allocated byte string) all data that has been written to the given string output port so far and fills in \*len with the length of the string in bytes (not including the nul terminator).

Creates a pair of ports, setting \*read and \*write; data written to \*write can be read back out of \*read. The pipe can store arbitrarily many unread characters,

Like scheme\_pipe is *limit* is 0. If *limit* is positive, creates a pipe that stores at most *limit* unread characters, blocking writes when the pipe is full.

```
Scheme_Input_Port* scheme_input_port_record(Scheme_Object* port)
```

Returns the input-port record for *port*, which may be either a raw-port object with type scheme\_input\_port\_type or a structure with the prop:input-port property.

```
Scheme_Output_Port*
scheme_output_port_record(Scheme_Object* port)
```

Returns the output-port record for *port*, which may be either a raw-port object with type scheme\_output\_port\_type or a structure with the prop:output-port property.

```
int scheme_file_exists(char* name)
```

Returns 1 if a file by the given name exists, 0 otherwise. If *name* specifies a directory, FALSE is returned. The *name* should be already expanded.

```
int scheme_directory_exists(char* name)
```

Returns 1 if a directory by the given name exists, 0 otherwise. The *name* should be already expanded.

Cleanses the pathname *name* (see cleanse-path) and resolves relative paths with respect to the current directory parameter. The *len* argument is the length of the input string; if it is -1, the string is assumed to be null-terminated. The *where* argument is used to raise an exception if there is an error in the filename; if this is NULL, an error is not reported and NULL is returned instead. If *expanded* is not NULL, \**expanded* is set to 1 if some expansion takes place, or 0 if the input name is simply returned.

If *guards* is not 0, then scheme\_security\_check\_file (see §15 "Security Guards") is called with *name*, *where*, and *checks* (which implies that *where* should never be NULL unless *guards* is 0). Normally, *guards* should be SCHEME\_GUARD\_FILE\_EXISTS at a minimum. Note that a failed access check will result in an exception.

Like scheme\_expand\_string, but given a *name* that can be a character string or a path value.

```
Scheme_Object* scheme_char_string_to_path(Scheme_Object* s)
```

Converts a Scheme character string into a Scheme path value.

```
Scheme_Object* scheme_path_to_char_string(Scheme_Object* s)
```

Converts a Scheme path value into a Scheme character string.

```
Scheme_Object* scheme_make_path(char* bytes)
```

Makes a path value given a byte string. The bytes string is copied.

```
Scheme_Object* scheme_make_path_without_copying(char* bytes)
```

Like scheme\_make\_path, but the string is not copied.

Makes a path whose byte form has size *len*. A copy of *bytes* is made if *copy* is not 0. The string *bytes* should contain *len* bytes, and if *copy* is zero, *bytes* must have a nul terminator in addition. If *len* is negative, then the nul-terminated length of *bytes* is used for the length.

```
Scheme_Object* scheme_make_sized_offset_path(char* bytes,
long d,
long len,
int copy)
```

Like scheme\_make\_sized\_path, except the *len* bytes start from position d in *bytes*. If d is non-zero, then copy must be non-zero.

```
char* scheme_build_mac_filename(FSSpec* spec, int isdir)
```

Mac OS X only: Converts an FSSpec record (defined by Mac OS X) into a pathname string. If *spec* contains only directory information (via the vRefNum and parID fields), *isdir* should be 1, otherwise it should be 0.

Mac OS X only: Converts a pathname into an FSSpec record (defined by Mac OS X), returning 1 if successful and 0 otherwise. If *type* is not NULL and *filename* is a file that exists, *type* is filled with the file's four-character Mac OS X type. If *type* is not NULL and *filename* is not a file that exists, *type* is filled with 0.

Gets the current working directory according to the operating system. This is separate from MzScheme's current directory parameter.

The directory path is written into *buf*, of length *buflen*, if it fits. Otherwise, a new (collectable) string is allocated for the directory path. If *actlen* is not NULL, \**actlen* is set to

the length of the current directory path. If *noexn* is no 0, then an exception is raised if the operation fails.

Sets the current working directory according to the operating system. This is separate from MzScheme's current directory parameter.

If *noexn* is not 0, then an exception is raised if the operation fails.

Creates a string like MzScheme's format procedure, using the format string *format* (of length *flen*) and the extra arguments specified in *argc* and *argv*. If *rlen* is not NULL, \**rlen* is filled with the length of the resulting string.

Writes to the current output port like MzScheme's printf procedure, using the format string *format* (of length *flen*) and the extra arguments specified in *argc* and *argv*.

Like scheme\_format, but takes a UTF-8-encoding byte string.

Like scheme\_printf, but takes a UTF-8-encoding byte string.

int scheme\_close\_should\_force\_port\_closed()

This function must be called by the close function for a port created with  ${\tt scheme\_make\_output\_port}.$ 

### 14 Structures

A new Scheme structure type is created with scheme\_make\_struct\_type. This creates the structure type, but does not generate the constructor, etc. procedures. The scheme\_make\_struct\_values function takes a structure type and creates these procedures. The scheme\_make\_struct\_names function generates the standard structure procedures names given the structure type's name. Instances of a structure type are created with scheme\_make\_struct\_instance and the function scheme\_is\_struct\_instance tests a structure's type. The scheme\_struct\_ref and scheme\_struct\_set functions access or modify a field of a structure.

The the structure procedure values and names generated by scheme\_make\_struct\_values and scheme\_make\_struct\_names can be restricted by passing any combination of these flags:

- SCHEME\_STRUCT\_NO\_TYPE the structure type value/name is not returned.
- SCHEME\_STRUCT\_NO\_CONSTR the constructor procedure value/name is not returned.
- SCHEME\_STRUCT\_NO\_PRED— the predicate procedure value/name is not returned.
- SCHEME\_STRUCT\_NO\_GET the selector procedure values/names are not returned.
- SCHEME\_STRUCT\_NO\_SET the mutator procedure values/names are not returned.
- SCHEME\_STRUCT\_GEN\_GET the field-independent selector procedure value/name is returned.
- SCHEME\_STRUCT\_GEN\_SET the field-independent mutator procedure value/name is returned.

When all values or names are returned, they are returned as an array with the following order: structure type, constructor, predicate, first selector, first mutator, second selector, etc., field-independent select, field-independent mutator. When particular values/names are omitted, the array is compressed accordingly.

Creates and returns a new structure type. The <code>base\_name</code> argument is used as the name of the new structure type; it must be a symbol. The <code>super\_type</code> argument should be NULL or an existing structure type to use as the super-type. The <code>inspector</code> argument should be NULL or an inspector to manage the type. The <code>num\_init\_fields</code> argument specifies the number of fields for instances of this structure type that have corresponding constructor arguments. (If a super-type is used, this is the number of additional fields, rather than the total number.) The <code>num\_auto\_fields</code> argument specifies the number of additional fields that have no corresponding constructor arguments, and they are initialized to <code>auto\_val</code>. The <code>properties</code> argument is a list of property-value pairs. The <code>guard</code> argument is either NULL or a procedure to use as a constructor guard.

Creates and returns an array of standard structure value name symbols. The *base\_name* argument is used as the name of the structure type; it should be the same symbol passed to the associated call to scheme\_make\_struct\_type. The *field\_names* argument is a (Scheme) list of field name symbols. The *flags* argument specifies which names should be generated, and if *count\_out* is not NULL, *count\_out* is filled with the number of names returned in the array.

Creates and returns an array of the standard structure value and procedure values for <code>struct\_type</code>. The <code>struct\_type</code> argument must be a structure type value created by <code>scheme\_make\_struct\_type</code>. The <code>names</code> procedure must be an array of name symbols, generally the array returned by <code>scheme\_make\_struct\_names</code>. The <code>count</code> argument specifies the length of the <code>names</code> array (and therefore the number of expected return values) and the <code>flags</code> argument specifies which values should be generated.

Creates an instance of the structure type *struct\_type*. The *argc* and *argv* arguments provide

the field values for the new instance.

Returns 1 if *v* is an instance of *struct\_type* or 0 otherwise.

Returns the nth field (counting from 0) in the structure s.

Sets the nth field (counting from 0) in the structure s to v.

# 15 Security Guards

Before a primitive procedure accesses the filesystem or creates a network connection, it should first consult the current security guard to determine whether such access is allowed for the current thread.

File access is normally preceded by a call to scheme\_expand\_filename, which accepts flags to indicate the kind of filesystem access needed, so that the security guard is consulted automatically.

An explicit filesystem-access check can be made by calling scheme\_security\_check\_file. Similarly, an explicit network-access check is performed by calling scheme\_security\_check\_network.

Consults the current security manager to determine whether access is allowed to *filename*. The *guards* argument should be a bitwise combination of the following:

- SCHEME\_GUARD\_FILE\_READ
- SCHEME\_GUARD\_FILE\_WRITE
- SCHEME\_GUARD\_FILE\_EXECUTE
- SCHEME\_GUARD\_FILE\_DELETE
- SCHEME\_GUARD\_FILE\_EXISTS (do not combine with other values)

The *filename* argument can be NULL (in which case #f is sent to the security manager's procedure), and *guards* should be SCHEME\_GUARD\_FILE\_EXISTS in that case.

If access is denied, an exception is raised.

Consults the current security manager to determine whether access is allowed for creating a client connection to *host* on port number *portno*. If *host* is NULL, the security manager is consulted for creating a server at port number *portno*.

If access is denied, an exception is raised.

### 16 Custodians

When an extension allocates resources that must be explicitly freed (in the same way that a port must be explicitly closed), a Scheme object associated with the resource should be placed into the management of the current custodian with scheme\_add\_managed.

Before allocating the resource, call scheme\_custodian\_check\_available to ensure that the relevant custodian is not already shut down. If it is, scheme\_custodian\_check\_available will raise an exception. If the custodian is shut down when scheme\_add\_managed is called, the close function provided to scheme\_add\_managed will be called immediately, and no exception will be reported.

Scheme\_Custodian\* scheme\_make\_custodian(Scheme\_Custodian\* m)

Creates a new custodian as a subordinate of m. If m is NULL, then the main custodian is used as the new custodian's supervisor. Do not use NULL for m unless you intend to create an especially privileged custodian.

Places the value o into the management of the custodian m. If m is NULL, the current custodian is used.

The f function is called by the custodian if it is ever asked to "shutdown" its values; o and data are passed on to f, which has the type

If *strong* is non-zero, then the newly managed value will be remembered until either the custodian shuts it down or scheme\_remove\_managed is called. If *strong* is zero, the value is allowed to be garbaged collected (and automatically removed from the custodian).

The return value from scheme\_add\_managed can be used to refer to the value's custodian later in a call to scheme\_remove\_managed. A value can be registered with at most one custodian.

If m (or the current custodian if m is NULL) is shut down, then f is called immediately, and the result is NULL.

```
void
```

Removes o from the management of its custodian. The mref argument must be a value returned by scheme\_add\_managed or NULL.

```
void scheme\_close\_managed(Scheme\_Custodian* m)
```

Instructs the custodian m to shutdown all of its managed values.

```
\verb|void scheme_add_atexit_closer(Scheme_Exit_Closer_Func|f)|\\
```

Installs a function to be called on each custodian-registered item and its closer when MzScheme is about to exit. The registered function has the type

where d is the second argument for f.

## 17 Miscellaneous Utilities

The MZSCHEME\_VERSION preprocessor macro is defined as a string describing the version of Scheme. The MZSCHEME\_VERSION\_MAJOR and MZSCHEME\_VERSION\_MINOR macros are defined as the major and minor version numbers, respectively.

```
int scheme_eq(Scheme_Object* obj1,
               Scheme_Object* obj2)
Returns 1 if the Scheme values are eq?.
int scheme_eqv(Scheme_Object* obj1,
                Scheme_Object* obj2)
Returns 1 if the Scheme values are eqv?.
int scheme_equal(Scheme_Object* objl,
                   Scheme_Object* obj2)
Returns 1 if the Scheme values are equal?.
int scheme_recur_equal(Scheme_Object* obj1,
                         Scheme_Object* obj2,
                         void* cycle_data)
Like scheme_equal, but accepts an extra value for cycle tracking. This procedure is meant
to be called by a procedure installed with scheme_set_type_equality.
Returns 1 if the Scheme values are equal?.}
long scheme_equal_hash_key(Scheme_Object* obj)
Returns the primary equal?-hash key for \var{obj}.
long scheme_equal_hash_key2(Scheme_Object* obj)
Returns the secondary equal?-hash key for \var{obj}.
long scheme_recur_equal_hash_key(Scheme_Object* obj,
                                    void* cycle_data)
```

Like scheme\_equal\_hash\_key, but accepts an extra value for cycle tracking. This procedure is meant to be called by a hasing procedure installed with scheme\_set\_type\_equality.

Returns the primary equal?-hash key for \var{obj}.}

Like scheme\_equal\_hash\_key2, but accepts an extra value for cycle tracking. This procedure is meant to be called by a secondary hashing procedure installed with scheme\_set\_type\_equality.

Creates and returns a list of length c with the elements elems.

```
int scheme_list_length(Scheme_Object* list)
```

Returns the length of the list. If *list* is not a proper list, then the last cdr counts as an item. If there is a cycle in *list* (involving only cdrs), this procedure will not terminate.

```
int scheme_proper_list_length(Scheme_Object* list)
```

Returns the length of the list, or -1 if it is not a proper list. If there is a cycle in *list* (involving only cdrs), this procedure returns -1.

```
Scheme_Object* scheme_car(Scheme_Object* pair)
```

Returns the car of the pair.

```
Scheme_Object* scheme_cdr(Scheme_Object* pair)
```

Returns the cdr of the pair.

```
Scheme_Object* scheme_cadr(Scheme_Object* pair)
```

Returns the cadr of the pair.

```
Scheme_Object* scheme_caddr(Scheme_Object* pair)
Returns the caddr of the pair.
Scheme_Object* scheme_vector_to_list(Scheme_Object* vec)
Creates a list with the same elements as the given vector.
Scheme_Object* scheme_list_to_vector(Scheme_Object* list)
Creates a vector with the same elements as the given list.
Scheme_Object* scheme_append(Scheme_Object* lstx,
                                Scheme_Object* lsty)
Non-destructively appends the given lists.
Scheme_Object* scheme_unbox(Scheme_Object* obj)
Returns the contents of the given box.
void scheme_set_box(Scheme_Object* b,
                      Scheme_Object* v)
Sets the contents of the given box.
Scheme_Object* scheme_dynamic_require(int argc,
                                          Scheme_Object** argv)
The same as dynamic-require. The argc argument must be 2, and argv contains the
arguments.
 Scheme_Object*
scheme_namespace_require(Scheme_Object* prim_req_spec)
The same as namespace-require.
```

Scheme\_Object\* scheme\_load(char\* file)

Loads the specified Scheme file, returning the value of the last expression loaded, or NULL if the load fails.

```
Scheme_Object* scheme_load_extension(char* filename)
```

Loads the specified Scheme extension file, returning the value provided by the extension's initialization function.

```
Scheme_Hash_Table* scheme_make_hash_table(int type)
```

Creates a hash table. The *type* argument must be either SCHEME\_hash\_ptr or SCHEME\_hash\_string, which determines how keys are compared (unless the hash and compare functions are modified in the hash table record; see below). A SCHEME\_hash\_ptr table hashes on a key's pointer address, while SCHEME\_hash\_string uses a key as a char\* and hashes on the null-terminated string content. Since a hash table created with SCHEME\_hash\_string (instead of SCHEME\_hash\_ptr) does not use a key as a Scheme value, it cannot be used from Scheme code.

Although the hash table interface uses the type Scheme\_Object\* for both keys and values, the table functions never inspect values, and they inspect keys only for SCHEME\_hash\_string hashing. Thus, the actual types of the values (and keys, for SCHEME\_hash\_ptr tables) can be anything.

The public portion of the Scheme\_Hash\_Table type is defined roughly as follows:

```
typedef struct Scheme_Hash_Table {
   Scheme_Object so; /* so.type == scheme_hash_table_type */
   /* ... */
   int size; /* size of keys and vals arrays */
   int count; /* number of mapped keys */
   Scheme_Object **keys;
   Scheme_Object **vals;
   void (*make_hash_indices)(void *v, long *h1, long *h2);
   int (*compare)(void *v1, void *v2);
   /* ... */
} Scheme_Hash_Table;
```

The make\_hash\_indices and compare function pointers can be set to arbitrary hashing and comparison functions (before any mapping is installed into the table). A hash function should fill h1 with a primary hash value and h2 with a secondary hash value; the values are for double-hashing, where the caller takes appropriate modulos. Either h1 or h2 can be NULL if the corresponding hash code is not needed.

To traverse the hash table content, iterate over *keys* and *vals* in parallel from 0 to size-1, and ignore *keys* where the corresponding *vals* entry is NULL. The count field indicates the

number of non-NULL values that will be encountered.

```
Scheme_Hash_Table* scheme_make_hash_table_equal()
```

Like scheme\_make\_hash\_table, except that keys are treated as Scheme values and hashed based on equal?\ instead of eq?.

Sets the current value for key in table to val. If val is NULL, the key is unmapped in table.

```
Scheme_Object* scheme_hash_get(Scheme_Hash_Table* table,
Scheme_Object* key)
```

Returns the current value for key in table, or NULL if key has no value.

Like make\_hash\_table, but bucket tables are somewhat more flexible, in that hash buckets are accessible and weak keys are supported. (They also consume more space than hash tables.)

The *type* argument must be either SCHEME\_hash\_ptr, SCHEME\_hash\_string, or SCHEME\_hash\_weak\_ptr. The first two are the same as for hash tables. The last is like SCHEME\_hash\_ptr, but the keys are weakly held.

The public portion of the Scheme\_Bucket\_Table type is defined roughly as follows:

```
typedef struct Scheme_Bucket_Table {
   Scheme_Object so; /* so.type == scheme_variable_type */
   /* ... */
   int size; /* size of buckets array */
   int count; /* number of buckets, >= number of mapped keys */
   Scheme_Bucket **buckets;
   void (*make_hash_indices)(void *v, long *h1, long *h2);
   int (*compare)(void *v1, void *v2);
   /* ... */
} Scheme_Bucket_Table;
```

The make\_hash\_indices and compare functions are used as for hash tables. Note that

SCHEME\_hash\_weak\_ptr supplied as the initial type makes keys weak even if the hash and comparison functions are changed.

See scheme\_bucket\_from\_table for information on buckets.

Sets the current value for key in table to val. If const is non-zero, the value for key must never be changed.

Sets the current value for key in table to val, but only if key is already mapped in the table.

Returns the current value for key in table, or NULL if key has no value.

Returns the bucket for key in table. The Scheme\_Bucket structure is defined as:

```
typedef struct Scheme_Bucket {
   Scheme_Object so; /* so.type == scheme_bucket_type */
   /* ... */
   void *key;
   void *val;
} Scheme_Bucket;
```

Setting *val* to NULL unmaps the bucket's key, and *key* can be NULL in that case as well. If the table holds keys weakly, then *key* points to a (weak) pointer to the actual key, and the weak pointer's value can be NULL.

```
long scheme_double_to_int(char* where, double d)
```

Returns a fixnum value for the given floating-point number d. If d is not an integer or if it is too large, then an error message is reported; name is used for error-reporting.

long scheme\_get\_milliseconds()

Returns the current "time" in milliseconds, just like current-milliseconds.

 ${\tt long \ scheme\_get\_process\_milliseconds()}$ 

Returns the current process "time" in milliseconds, just like current-process-milliseconds.

char\* scheme\_banner()

Returns the string that is used as the Scheme startup banner.

char\* scheme\_version()

Returns a string for the executing version of Scheme.

# 18 Flags and Hooks

The following flags and hooks are available when PLT Scheme is embedded:

- scheme\_exit This pointer can be set to a function that takes an integer argument and returns void; the function will be used as the default exit handler. The default is NULL.
- scheme\_make\_stdin, scheme\_make\_stdout, scheme\_make\_stderr, These pointers can be set to a function that takes no arguments and returns a Scheme port Scheme\_Object \* to be used as the starting standard input, output, and/or error port. The defaults are NULL. Setting the initial error port is particularly important for seeing unexpected error messages if stderr output goes nowhere.
- scheme\_console\_output This pointer can be set to a function that takes a string and a long string length; the function will be called to display internal MzScheme warnings and messages that possibly contain non-terminating nuls. The default is *NULL*.
- scheme\_check\_for\_break This points to a function of no arguments that returns an integer. It is used as the default user-break polling procedure in the main thread. A non-zero return value indicates a user break, and each time the function returns a non-zero value, it counts as a new break signal (though the break signal may be ignored if a previous signal is still pending). The default is NULL.
- scheme\_case\_sensitive If this flag is set to a non-zero value before scheme\_basic\_env is called, then MzScheme will not ignore capitalization for symbols and global variable names. The value of this flag should not change once it is set. The default is zero.
- scheme\_allow\_set\_undefined This flag determines the initial value of compile-allow-set!-undefined. The default is zero.
- scheme\_console\_printf This function pointer was left for backward compatibility. The default builds a string and calls scheme\_console\_output.

Initializes the current-library-collection-paths parameter using find-library-collection-paths. The *extra\_paths* argument is propagated to find-library-collection-paths.

| Indov                                   |  |
|---|--|
| Index                                   | continuations, 58  |
| "escheme.h",5                           | Cooperating with 3m, 30  |
| "libmzgc.a", 8                          | current working directory, 110   |
| "libmzgc.la", 8                         | Custodians, 118  |
| "libmzgc.so", 8                         | embedding MzScheme   |
| "libmzscheme.a", 8                      | Embedding MzScheme into a Program, 8   |
| "libmzscheme.la", 8                     | Enabling and Disabling Breaks, 61  |
| "libmzscheme.so", 8                     | Evaluation, 52   |
| "libmzscheme3m.a",12                    | Evaluation Functions, 54   |
| "libmzscheme3m.la",12                   | Exception Functions, 62  |
| "libmzscheme3m.so", 12                  | Exceptions and Escape Continuations, 58  |
| "mzdyn.o", 6                            | extending MzScheme, 5  |
| "mzdyn.obj",6                           | Flags and Hooks  |
| "mzdyn3m.o", 8                          | garbage collection   |
| "mzdyn3m.obj", 8                        | GC_register_traversers, 44 Global Constants, 19  |
| "scheme.h",9                            | globals, in extension code, 30   |
| #%variable-reference, 46                | globals, 46  |
| 3m, 8                                   | Guiding mzcxform, 37   |
| cc, 6                                   | header files   |
| cgc, 6                                  | Inside: PLT Scheme C API   |
| ld, 6                                   | Integers, 14   |
| xform, 8                                | Integers, 14 Integers and 14 I |
| 3m, 5                                   | Local Pointers   |
| 3m Embedding, 11                        | Local Pointers and mzcxform, 36  |
| 3m Extensions, 7                        | memory   |
| _scheme_apply, 55                       | Memory Allocation, 29  |
| _scheme_apply_multi, 55                 | Memory Functions, 39   |
| _scheme_eval_compiled, 54               | Miscellaneous Utilities, 120   |
| _scheme_eval_compiled_multi, 54         | module, 46   |
| allocation                              | Multiple Values, 53  |
| allocation, 10                          | mzc, 6   |
| Allowing Thread Switches, 68            | MZCONFIG_ALLOW_SET_UNDEFINED, 82   |
| arity,49                                | MZCONFIG_CAN_READ_BOX, 81  |
| Bignums, Rationals, and Complex Numbers | MZCONFIG_CAN_READ_COMPILED, 81   |
| Blocking the Current Thread, 68         | MZCONFIG_CAN_READ_GRAPH, 81  |
| Callbacks for Blocked Threads           | MZCONFIG_CAN_READ_PIPE_QUOTE, 81   |
| CGC, 5                                  | MZCONFIG_CASE_SENS, 82   |
| CGC Embedding, 8                        | MZCONFIG_COLLECTION_PATHS, 82  |
| CGC Extensions, 5                       | MZCONFIG_CURLY_BRACES_ARE_PARENS, 82   |
| CGC versus 3m, 5                        | MZCONFIG_CUSTODIAN, 82   |
| Continuation Marks, 85                  | MZCONFIG_ENV, 81   |
|   |  |

MZCONFIG\_ERROR\_DISPLAY\_HANDLER, 81 scheme\_add\_swap\_callback, 80 MZCONFIG\_ERROR\_PORT, 81 scheme\_add\_swap\_out\_callback, 80 MZCONFIG\_ERROR\_PRINT\_VALUE\_HANDLER, scheme\_add\_to\_table, 125 81 scheme\_alloc\_byte\_string, 22 MZCONFIG\_ERROR\_PRINT\_WIDTH, 82 scheme\_alloc\_char\_string, 24 MZCONFIG\_EVAL\_HANDLER, 81 scheme\_allow\_set\_undefined, 127 MZCONFIG\_EXIT\_HANDLER, 81 scheme\_append, 122 MZCONFIG\_INIT\_EXN\_HANDLER, 81 scheme\_append\_byte\_string, 23 MZCONFIG\_INPUT\_PORT, 81 scheme\_append\_char\_string, 24 MZCONFIG\_LOAD\_DIRECTORY, 82 scheme\_apply, 55 MZCONFIG\_LOAD\_EXTENSION\_HANDLER, 82 scheme\_apply\_multi, 55 MZCONFIG\_LOAD\_HANDLER, 81 scheme\_apply\_to\_list, 55 MZCONFIG\_OUTPUT\_PORT, 81 scheme\_banner, 126 MZCONFIG\_PORT\_PRINT\_HANDLER, 82 scheme\_basic\_env, 54 MZCONFIG\_PRINT\_BOX, 82 scheme\_bignum\_from\_double, 92 MZCONFIG\_PRINT\_GRAPH, 81 scheme\_bignum\_from\_float, 92 MZCONFIG\_PRINT\_HANDLER, 81 scheme\_bignum\_normalize, 92 MZCONFIG\_PRINT\_STRUCT, 82  $scheme\_bignum\_to\_double, 91$ MZCONFIG\_PROMPT\_READ\_HANDLER, 81 scheme\_bignum\_to\_float, 92 MZCONFIG\_SQUARE\_BRACKETS\_ARE\_PARENS, scheme\_bignum\_to\_string, 92 scheme\_bignum\_type, 16 MZCONFIG\_USE\_COMPILED\_KIND, 82 scheme\_block\_until, 75 MzScheme and Threads, 14  ${\tt scheme\_block\_until\_enable\_break,76}$ MzScheme, Unicode, Characters, scheme\_block\_until\_unless, 76 Strings, 14 scheme\_bool\_type, 16 Namespaces and Modules scheme\_box, 26 Overview scheme\_box\_type, 17 Parameterizations scheme\_break\_thread, 75 Ports and the Filesystem, 95 scheme\_break\_waiting, 75 Procedures, 49 scheme\_bucket\_from\_table, 125 scheme\_add\_atexit\_closer scheme\_bucket\_table\_type, 17 scheme\_add\_evt, 78 scheme\_build\_list, 121 scheme\_add\_evt\_through\_sema, 78 scheme\_build\_mac\_filename, 110 scheme\_add\_fd\_eventmask, 77 scheme\_builtin\_value, 47 scheme\_add\_fd\_handle,77 scheme\_byte\_ready, 100 scheme\_add\_finalizer, 43 scheme\_byte\_string\_to\_char\_string, scheme\_add\_finalizer\_once, 43 scheme\_add\_global, 46 scheme\_byte\_string\_to\_char\_string\_locale, scheme\_add\_global\_symbol, 46 scheme\_add\_managed, 118 scheme\_byte\_string\_type, 16 scheme\_add\_scheme\_finalizer, 43 scheme\_caddr, 122 scheme\_add\_scheme\_finalizer\_once, 43 scheme\_cadr, 121

```
scheme_call_enable_break, 79
                                       scheme_detach_multiple_array, 57
scheme_calloc, 39
                                       scheme_directory_exists, 108
scheme_car, 121
                                       scheme_display, 95
                                       scheme_display_to_string, 97
scheme_case_closure_type, 18
scheme_case_sensitive, 127
                                       scheme_display_to_string_w_max, 97
scheme_cdr. 121
                                       scheme_display_w_max, 96
scheme_change_in_table, 125
                                       scheme_dont_gc_ptr, 44
scheme_char_ready, 100
                                       scheme_double_to_int, 125
scheme_char_string_to_byte_string,
                                       scheme_double_type, 16
                                       scheme_dynamic_require, 122
scheme_char_string_to_byte_string_localeneme_dynamic_wind, 65
                                       scheme_end_atomic, 79
scheme_char_string_to_path, 109
                                       scheme_end_atomic_no_swap, 80
scheme_char_string_type, 16
                                       scheme_eof, 19
scheme_char_type, 16
                                       scheme_eq, 120
scheme_check_for_break, 127
                                       scheme_equal, 120
scheme_check_proc_arity, 65
                                       scheme_equal_hash_key, 120
scheme_check_threads, 76
                                       scheme_equal_hash_key2, 120
scheme_clear_escape, 65
                                       scheme_eqv, 120
scheme_close_input_port, 100
                                       scheme_escaping_cont_type, 18
scheme_close_managed, 119
                                       scheme_eval, 54
scheme_close_output_port, 101
                                       scheme_eval_compiled, 54
{\tt scheme\_close\_should\_force\_port\_closed\_scheme\_eval\_compiled\_multi, 54}
  112
                                       scheme_eval_string, 55
scheme_closed_prim_type, 18
                                       scheme_eval_string_all, 56
scheme_collect_garbage, 44
                                       scheme_eval_string_multi, 56
scheme_compile, 56
                                       SCHEME_EXACT_INTEGERP, 18
scheme_compiled_closure_type, 18
                                       SCHEME_EXACT_REALP. 18
scheme_complex_imaginary_part, 94
                                       scheme_exit, 127
scheme_complex_izi_type, 16
                                       scheme_expand, 57
scheme_complex_real_part, 93
                                       scheme_expand_filename, 109
scheme_complex_type, 16
                                       scheme_expand_string_filename, 109
scheme_console_output, 127
                                       scheme_extend_config, 83
scheme_console_printf, 127
                                       scheme_false, 19
scheme_cont_type, 18
                                       scheme_file_exists, 108
scheme_count_lines, 100
                                       scheme\_finish\_primitive\_module, 48
scheme_cpointer_type, 18
                                       scheme_float_type, 16
scheme_current_continuation_marks,
                                       SCHEME_FLOATP, 18
 66
                                       scheme_flush_output, 97
scheme_custodian_check_available,
                                       scheme_format, 111
                                       scheme_format_utf8, 111
scheme_debug_print, 97
                                       scheme_free_immobile_box, 40
```

```
scheme_gc_ptr_ok, 44
                                       scheme_list_to_vector, 122
scheme_get_byte, 98
                                       scheme_load, 122
scheme_get_byte_string, 98
                                       scheme_load_extension, 123
scheme_get_byte_string_output, 107
                                       scheme_lookup_global, 47
scheme_get_bytes, 99
                                       scheme_lookup_in_table, 125
scheme_get_char_string, 99
                                       scheme_mac_path_to_spec, 110
scheme_get_env, 48
                                       scheme_main_setup, 40
scheme_get_fdset, 77
                                       scheme_make_args_string, 64
scheme_get_int_val, 21
                                       scheme_make_ascii_character, 20
{\tt scheme\_get\_long\_long\_val}, 21
                                       scheme_make_bignum, 91
scheme_get_milliseconds, 126
                                       scheme_make_bignum_from_unsigned, 91
scheme_get_param, 82
                                       scheme_make_bucket_table, 124
scheme_get_port_fd, 101
                                       scheme_make_byte_string, 22
scheme_get_port_file_descriptor, 101
                                       scheme_make_byte_string_output_port,
{\tt scheme\_get\_process\_milliseconds}, 126
scheme_get_sized_byte_string_output,
                                       scheme_make_byte_string_without_copying,
                                         22
                                       scheme_make_char, 19
scheme_get_thread_param, 83
                                       scheme_make_char_or_null, 20
scheme_get_unsigned_int_val, 21
                                       scheme_make_char_string, 23
scheme_get_unsigned_long_long_val,
                                       scheme_make_char_string_without_copying,
scheme_getc, 98
                                       scheme_make_character, 20
scheme_global_bucket, 47
scheme_hash_get, 124
                                       scheme_make_closed_prim, 51
                                       scheme_make_closed_prim_w_arity, 51
scheme_hash_set, 124
scheme_hash_table_type, 17
                                       scheme_make_complex, 93
                                       scheme_make_cptr, 26
scheme_inherit_cells, 83
                                       scheme_make_custodian, 118
scheme_init_collection_paths, 127
                                       scheme_make_double, 21
scheme_input_port_record, 108
                                       scheme_make_exact_symbol, 25
scheme_input_port_type, 17
                                       scheme_make_fd_input_port, 107
scheme_install_config,83
                                       scheme_make_file_input_port, 106
scheme_integer_type, 16
                                       scheme_make_file_output_port, 107
scheme_intern_exact_char_keyword, 26
                                       scheme_make_float, 21
scheme_intern_exact_char_symbol, 25
                                       scheme_make_folding_prim, 50
scheme_intern_exact_keyword, 25
                                       scheme_make_hash_table, 123
scheme_intern_exact_symbol, 25
scheme_intern_symbol, 25
                                       scheme_make_hash_table_equal, 124
                                       scheme_make_input_port, 101
scheme_is_exact, 91
                                       scheme_make_integer, 20
scheme_is_inexact, 91
                                       scheme_make_integer_value, 20
scheme_is_struct_instance, 115
                                       scheme_make_integer_value_from_long_halves,
scheme_keyword_type, 17
                                         20
scheme_list_length, 121
```

```
scheme_make_integer_value_from_long_longheme_make_symbol, 25
                                       scheme_make_thread_cell, 79
scheme_make_integer_value_from_unsign@cheme_make_type, 26
                                       scheme_make_utf8_string, 23
scheme_make_integer_value_from_unsigned_heng_halvesctor. 26
 21
                                       scheme_make_weak_box. 26
scheme_make_integer_value_from_unsigned_long_long_progress, 78
 20
                                       scheme_malloc, 39
scheme_make_locale_string, 23
                                       scheme_malloc_allow_interior, 39
scheme_make_named_file_input_port,
                                       scheme_malloc_atomic, 39
                                       scheme_malloc_eternal, 39
scheme_make_namespace, 55
                                       scheme_malloc_fail_ok, 40
scheme_make_offset_cptr, 26
                                       scheme_malloc_immobile_box, 40
scheme_make_output_port, 104
                                       scheme_malloc_tagged, 39
scheme_make_pair, 22
                                       scheme_malloc_uncollectable, 39
scheme_make_path, 109
                                       scheme_module_bucket, 47
scheme_make_path_without_copying,
                                       scheme_namespace_require, 122
                                       scheme_namespace_type, 18
scheme_make_port_type, 101
                                       scheme_native_closure_type, 18
scheme_make_prim, 50
                                       scheme_need_wakeup, 100
scheme_make_prim_closure_w_arity, 50
                                       scheme_new_param, 83
scheme_make_prim_w_arity, 49
                                       scheme_null, 19
scheme_make_provided_string, 64
                                       SCHEME_NUMBERP, 18
scheme_make_rational, 92
                                       scheme_open_input_file, 106
scheme_make_sema, 74
                                       scheme_open_output_file, 107
scheme_make_sized_byte_string, 22
                                       scheme_os_getcwd, 110
scheme_make_sized_char_string, 24
                                       scheme_os_setcwd, 111
scheme_make_sized_offset_byte_string,
                                       scheme_output_port_record, 108
 22
                                       scheme_output_port_type, 17
scheme_make_sized_offset_char_string,
                                       scheme_pair_type, 17
 24
                                       scheme_param_config, 84
scheme_make_sized_offset_path, 110
                                       scheme_path_to_char_string, 109
scheme_make_sized_offset_utf8_string,
                                       scheme_path_type, 16
 23
                                       scheme_peek_byte, 98
scheme_make_sized_path, 110
                                       scheme_peek_byte_skip, 98
scheme_make_sized_utf8_string, 23
                                       scheme_peekc, 98
scheme_make_stderr, 127
                                       scheme_peekc_skip, 98
scheme_make_stdin, 127
                                       scheme_pipe, 108
scheme_make_stdout, 127
                                       scheme\_pipe\_with\_limit, 108
scheme_make_struct_instance, 114
                                       scheme_pop_break_enable, 66
scheme_make_struct_names, 114
                                       scheme_pop_continuation_frame, 85
scheme_make_struct_type, 113
                                       scheme_post_sema, 74
scheme_make_struct_values, 114
```

scheme\_prim\_type, 18 scheme\_set\_stack\_base, 41 scheme\_primitive\_module, 48 scheme\_set\_stack\_bounds, 41 scheme\_print\_bytes, 27 scheme\_set\_thread\_param, 83 scheme\_print\_string, 27 scheme\_set\_type\_equality, 27 scheme\_printf, 111 scheme\_set\_type\_printer, 27 scheme\_printf\_utf8, 111 scheme\_signal\_error, 62 scheme\_proper\_list\_length, 121 scheme\_signal\_received, 76 scheme\_push\_break\_enable, 66 scheme\_strdup, 39 scheme\_push\_continuation\_frame, 85 scheme\_strdup\_eternal, 40 scheme\_put\_byte\_string, 96 scheme\_struct\_property\_type, 17 scheme\_put\_char\_string, 97 scheme\_struct\_ref, 115 scheme\_raise\_exn, 63 scheme\_struct\_set, 115 scheme\_rational\_denominator, 93 scheme\_struct\_type\_type, 17 scheme\_rational\_from\_double, 93 scheme\_structure\_type, 17 scheme\_rational\_from\_float, 93 scheme\_subtract\_finalizer, 44 scheme\_rational\_numerator, 93 scheme\_swap\_thread, 75 scheme\_rational\_to\_double, 93 scheme\_symbol\_type, 17 scheme\_rational\_to\_float, 93 scheme\_tail\_apply, 56 scheme\_rational\_type, 16 scheme\_tail\_apply\_no\_copy, 56 scheme\_read, 95 scheme\_tail\_apply\_to\_list, 56 scheme\_read\_bignum, 92 scheme\_tell, 100 scheme\_read\_bignum\_bytes, 92 scheme\_tell\_line, 100 scheme\_real\_to\_double, 22 scheme\_thread, 73 SCHEME\_REALP, 18 scheme\_thread\_block, 74 scheme\_thread\_block\_enable\_break, 75 scheme\_recur\_equal, 120 scheme\_recur\_equal\_hash\_key, 120 scheme\_thread\_type, 17 scheme\_recur\_equal\_hash\_key2, 121 scheme\_thread\_w\_details, 74 scheme\_register\_extension\_global, 40 scheme\_tls\_allocate, 79 scheme\_register\_finalizer, 42 scheme\_tls\_get, 79 scheme\_register\_parameter, 84 scheme\_tls\_set, 79 scheme\_register\_static, 42 scheme\_true, 19 scheme\_remove\_all\_finalization, 44 scheme\_ucs4\_to\_utf16,89 scheme\_security\_check\_file, 116 scheme\_unbound\_global, 64 scheme\_security\_check\_network, 116 scheme\_unbox, 122 scheme\_sema\_type, 17 scheme\_undefined, 19 scheme\_set\_box. 122 scheme\_ungetc, 99 scheme\_set\_can\_break, 65 scheme\_utf16\_to\_ucs4,90 scheme\_set\_collects\_path, 127 scheme\_utf8\_decode, 86  ${\tt scheme\_set\_cont\_mark},\,85$ scheme\_utf8\_decode\_all, 87  $scheme\_set\_file\_position, 100$ scheme\_utf8\_decode\_as\_prefix, 87 scheme\_set\_global\_bucket, 47 scheme\_utf8\_decode\_count, 88 scheme\_set\_param, 82 scheme\_utf8\_decode\_prefix, 87

scheme\_utf8\_decode\_to\_buffer, 88 scheme\_utf8\_decode\_to\_buffer\_len, 88 scheme\_utf8\_encode, 88 scheme\_utf8\_encode\_all, 89 scheme\_utf8\_encode\_to\_buffer, 89 scheme\_utf8\_encode\_to\_buffer\_len, 89 scheme\_values, 57 scheme\_vector\_to\_list, 122 scheme\_vector\_type, 17 scheme\_version, 126 scheme\_void, 19 scheme\_wait\_sema, 74 scheme\_wake\_up, 77 scheme\_warning, 63 scheme\_weak\_box\_type, 18 scheme\_weak\_reference, 42 scheme\_weak\_reference\_indirect, 42  $scheme\_write, 95$  $scheme\_write\_byte\_string, 96$ scheme\_write\_char\_string, 96 scheme\_write\_to\_string, 97 scheme\_write\_to\_string\_w\_max, 97 scheme\_write\_w\_max, 95 scheme\_wrong\_count, 63 scheme\_wrong\_return\_arity, 64 scheme\_wrong\_type, 63 Security Guards, 116 Sleeping by Embedded Scheme, 72 Standard Types, 16 String Encodings, 86 strings, conversion to C, 16 strings, conversion to C, 16 Strings, 19 Structures, 113 sync, 69 **Tagged Objects** Tail Evaluation, 52 tail recursion, 52 Temporarily Catching Error Escapes, 59 Thread Functions, 73 Threads, 67

Threads in Embedded Scheme with Event Loops, 69 Top-level Evaluation Functions, 52 types, creating, 15 Value Functions Values and Types, 15 Writing MzScheme Extensions